

```
Db 21 NLDVBPVIFVREDAAFGQTVVQFGSRLVVGAPLEAVVQNGRLYDCAPATGMCQPIV 80
Qy 62 LQVPVAVNMSLGLSAAATTSPPQLLACGPTVHTQCTSENTYVKGCLFLFGSNLRQOPKF 121
Db 81 LRSPLRAVNSLGLSVTATNNAQLACGPTAQRACVQKMYAGKSCILLGSSL-QFIQAV 139
Qy 122 PEARLQCPQSDSDIAFLIDSGSIIIPHDPRMKWSTVMEQLKSKTLFSLMOYSEFR 181
Db 140 PASMECPFROEMDIAFLIDSGSINORDFAQMDFAQMDFAQMDFAQMDFAQMDFAQMD 199
Qy 182 IHTTFKFNPNPRSLIKPITOLLGRTHATGLRKVVRELFINITNGARKNAFKILFLT 241
Db 200 THFTTFKFNLPQSLVDPIVOLQGLTYTATGIRTWELFHSKNGRSACKILLVIT 259
Qy 242 DGBKFGDPLGYEDVPELDREGVIRVIGVGDAPRSEKSRQELNTVASKPRDRHVQINN 301
Db 260 DGQKRDPLEYSDVIPAADKAGIRVAIGVGDAPRSEKSRQELNTVASKPRDRHVQINN 319
Qy 302 FEALKTIONLREKI FAIEGTQCTSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAGG 361
Db 320 FAALRSIQRLQOEKI FAIEGTQCTSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAGG 379
Qy 362 VFYLSKSKSTPNTMTVDSMDNDAYLGAAAAIILNRVOSLVGLGAPRYQHIGLVAMFRQ 421
Db 380 AFLYPNTPTPTFNMSENQVMDRSDYLGSTAVAFKGVHSLILGAPRHQHTGKVIPTQ 439
Qy 422 NTGMWESNANVKTQICAYFGASICSVDVDSNGSTDIAVLIGAPHYHQTGQGVSVCLP 481
Db 440 EAHWRPKSEVRGTQIGSYFGASLCSVDVDRDGDLDVLIGAPHYHQTGQGVSVFPVP 499
Qy 482 RQQRARQCDVAVLYGEOGPWGRFGAALTIVLGVNGDKLTDAVGAPEEDNRGAVYLFH 541
Db 500 -GVGRQWQCEATLHGXQHPWRGFGVALTIVLGVNGDNLADVAIGAPGESRGAIVYFH 558
Qy 542 GTSOGSISPHSQRIACSKSLPRLQYFGOSLSCGODLTMLGLVLTGACQHVILLRSQP 601
Db 559 GASLLEIMPSPSRVTSLSRLQYFGOSLSCGODLTMLGLVLTGACQHVILLRSQP 618
Qy 602 VLRVKALMESNPREVAZNFECNDQVVKGEAGVRVCLHVQKSTRDLREGQTQSVVTV 661
Db 619 LKVELSIRAPMEVAKAVQCWERTPTVLEAGATVCLTVHKGSPDLL-GNVQGSVRY 676
Qy 662 DLALDSRPSRAVFNETHKSTRQTOVLGHTQCTETLKLQPLNCIEDPSPVILRLNS 721
Db 677 DLALDPGLISRAIFEDTKNCTLTKRKTGLGDHCETVKKLLPDCVEDAVSPILRLNS 736
Qy 722 LVGTPLAGNLRVLAEDAORLTALPFPFKNGNDNI CODLSITFSFMSLDCLVVG 781
Db 737 LVRSASP-RNLHPLVAVGSQDHITASLPFEKCKQELLCEGDLGIFNFSGLQVLVVG 795
Qy 782 PREFNVTVNRDGEDSVRTQVTPFFPLDLSYRKVSTLQNRORSRWLACESASSTVS 841
Db 796 SPELTIVTVVWNEGEDSYGLVAKFYYPAGLSYRRTVGTQ-QPHQYPLRLACEAPAAQED 854
Qy 842 GALKSTCSINHPIFPENSEVTENITPDVDSKASLGNKLLKANVTSENNNPKNTKTEFO 901
Db 855 --LSSSCSINHPIFREGAKTITMITFDVSYKAFGLGRLLLRKASSENKPTNKTAFO 912
Qy 902 LEPLVKAVYVMTVSHGYSKYLNPASENTSR-VMOHQYOVNSLXGORSIPSLVPLVPV 960
Db 913 LEPLVKAVYVMTVSHGYSKYLNPASENTSR-VMOHQYOVNSLXGORSIPSLVPLVPV 972
Qy 961 PLNQTVIWRDQVTFSENLS--TCHTKERLPSHSDLAELRKAPVNVCSLAVOCQICDI 1019
Db 973 LINGAVWD----VTLSSPAQGVCSVQKPPQNPQDFITQIRRVLDCCSLADLHFRCDI 1029
Qy 1020 PFFGQIBEFNATLKNLSFDWYIKTSHNHLIVSTABILPNDVSFTLLPQOGAFVRSQTE 1079
Db 1030 PSLDQDELDFILGNLSFGVSVQTLQEKVLLVSEABITFTSVYSLPQGEAPLRAQVE 1089
Qy 1080 TKVPEFVNPDLPLVGSVSGILLALLITAAVLKIGEFKRYQKMM 1126
Db 1090 TTLSEYVYVYFPLVAGSVGGLLLALLITVVLKLGFFFRQYKEML 1136
```

RESULT 7

Q9WTV4

```
ID Q9WTV4 PRELIMINARY; PRT: 1161 AA.
AC Q9WTV4;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Integrin alpha L.
GN ITGAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=DBA/2J; TISSUE=Spleen;
RA Ma R.Z., Teuscher C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065902; AAD25885.1; -.
DR HSP; P20701; ILPA.
DR MGD; MGI:96606; Itgal.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; EG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
DR PROSITE; PS00234; VWFA; 1.
SQ SEQUENCE 1161 AA; 128240 MW; 86B102F7B209B431 CRC64;
```

Query Match 25.9%; Score 1523.5; DB 11; Length 1161;

Best Local Similarity 33.9%; Pred. No. 4.5e-103;

Matches 396; Conservative 220; Mismatches 461; Indels 91; Gaps 37;

Qy 1 FNLDTENAMTFQENA-RGFGSQSVQIQGSRVVGVAQOEIVAAANQKSLVQCDYSTGSCBP 59

Db 24 YNLDTPTQSFIAQAGRHFGYQVLQEDG-VVVGAPGE---GNTGGLTHCRTSSSEFCQP 79

Qy 60 IRLQVPEAVNMSLGLSAAATTSPPQLLACGPTVHTQCTSENTYVKGCLFLFGSNLRQOPQ 119

Db 80 VSLH-GSNHSTKYLGMTLATDAAGKSLACDPGLSRTCDQNTVLSGLCYLFPQSLGPM 138

Qy 120 KPEALRGCPQSDSDIAFLIDSGSIIIPHDPRMKWSTVMEQLKSKTLFSLMOYSEE 179

Db 139 QNRPAYQECMKGVLDVLFVFDGSQLDRKDFKILEFMKDVNRKJLSTNTSYQFAAVQFSTD 198

Qy 180 FRTHFTFKCF-ONNPNPRSLIKPITOLLGRTHATGLRKVVRELFINITNGARKNAFKILF 238

Db 199 CRTEFTFLVYKQKQNPVLLGSLVQVPMFLTTNTFRAINVVAHVFKESGARDPAKVLV 258

Qy 239 LFTDGEKF--GDLPGVEDVPELDREGVIRVIGVGDAPRSEKSRQELNTVASKPRDRHV 296

Db 259 IITDGEASDKENISAHD-----ITRYIIGIKGFVSVQKQTLHIFASEPEVEFV 309

Qy 297 FOINNFALKTNQLREKIPALEGTQCTSSSFHEMSQEGFSAAITNSGPLLSTVGSY 356

Db 310 KILDTTEKLDLPDLPORRIYAEIENRQDLTSFNNELSSGSLADLSKGHAVVGAVGAK 369

Qy 357 DMAGGVF-LYTSKSKSTFFINMTVDSMDNDAYLGAAAA-IILNRVQSLVGLGAPRYQHIG 414

Db 370 DMAGGFLDLREDLQGATFVQGBPLTSDVRGVLGYTVVAVMTSRSSRPLLAAGAPRYQHV 429

Qy 415 LVAMFR-QNTGMWESNANVKTQICAYFGASICSVDVDSNGSTDIAVLIGAPHYHQTG 472


```

Db 780 SSPARS-----CPLRMSSASLAVZTWLNSGEDAYVRLDLDPRGLSPKVEMLQ- 831
Qy 822 QRSORSRLACESASSTEVSGAL-KSTSCSINHPHPENSEVTFNITFDVDSKASLONKL 880
Db 832 --PHSRMPVSCHEL--TEGSLTKTKLCNVSSPIKAGQEVSLQVNFNTLMSWEDFV 887
Qy 881 LLKXNTVSEN-NMPTNKTETFOELPKVYAVYVMTVSHGVSTKYLNFTASENTRVMOHQ 939
Db 888 ELNGTVCENENSLOEDNSAATHIPVLYPNVILTXEQENSTLYISFTPKGPKTQQQHV 947
Qy 940 YQVSNLQORSPLSLVFLVRLNQTVINDRPQ-----VTFSEMLSS-----TCHKE-RLP 990
Db 948 YQV-----RIOPGAYDNMPT-LEALGVVERPHSEDLITYTWSVQTDPLVCHSEDLKRP 1001
Qy 991 SHSDFLAELRKAPVVCISAVCORIQCDIPFGIOBEFNATLKNLSFDVYIKTSENHLL 1050
Db 1002 S-----SEAPCLPGV--QFRCPVIF---RWEILQVGTVELSKHITAS-STLS 1045
Qy 1051 IVSTAELFNDVSFTLLPGQAFVRSQTEKVPFFVFNPLPIVGVSSVGLLLALITA 1110
Db 1046 LCSSLVSFNSSKHFLYGSKA-SEAQVLVKVDLHEKEMLVVVLGIGGLVLLFLFL 1104
Qy 1111 ALYKLGPFKEQYQDMW-SEGSPGAEP 1136
Db 1105 ALYKVGFFKGNLKEKMEADGGVNGSP 1131

RESULT 9
Q98TF1 PRELIMINARY; PRT; 1196 AA.
AC Q98TF1; 2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE C011-1.
GN C1A1.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Nakao M., Miura C., Fujiki K., Yano T.;
RT "Molecular cloning of a leukocyte integrin from the common carp.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB048536; BAB39134.1; -.
DR HSSP; P20701; 1LPA.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004190; P:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001969; A:protease AS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRIN.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50234; VWF_A; 1.
SQ SEQUENCE 1196 AA; 132477 MW; 9369C807B7DCA53B CRC64;

Query Match 23.7%; Score 1395; DB 13; Length 1196;
Best Local Similarity 31.2%; Pred. No. 1.5e-93;
Matches 371; Conservative 227; Mismatches 448; Indels 144; Gaps 39;
Qy 1 PNLDTENMTFQENARG-FCGSVVQLQ-GSR--VVVGAPQEIIVANQSGSLYQCDISTGS 56

```

```

Db 32 FNIDTEHPLRFNGAPEDPFGYSVYQTEFGNRKQIIVGAPLE---GNSGEMVSTACTAQOS 88
Qy 57 CDPRLQVP-----VRVNMISGLSLAATSPQQLACGPTVHTQTSENTYVVGKGLFLFOS 112
Db 89 CQ--BLQRPGSSEVRFFPGMSAAVSSAALTS-----CSFYPFHECDGNSYLVGVCYQFSS 140
Qy 113 NLROQPOKFPBALRGCPQEDSDIAFLIDGSGSIIIPHDPRMKKEWSTVMEQKLKSKTLFS 172
Db 141 SL-QAVSNPTAAYQECSEKENVLFLFDGSSSMKTVDFEMXNFKIKDINKKLSNSIKFA 199
Qy 173 LMQYSEEPRIHTPKFQNNPNRSLIKPI TOLLGRTHATGLRKVVRLEFN-INGARK 231
Db 200 AVQFSTVTRVTFDFNDYQSGSAEELMKE-TMKSLTNTTHKADIVYLKLLNLSMLSGADS 258
Qy 232 NAFKILPLLTGCKEKGDPGLGED--VIPELDREGVIRVVGVDAPFRSKSQEILNTVAS 289
Db 259 KAKKALVIITD---GDFSDDDYVNLKCDQNIIRYIIVGV-----KVDLIELTQLAS 309
Qy 290 KPRDRHVFOINNFEALKTIONQREKIPAEICTOTGSSSSPEHEMSQEGFSAIITSNGPL 349
Db 310 EPRKNTTFYIKDYSGLKGLLDNLQKKIYNIEGSVDAQGRDRQKLSQSGFSVYIOBESVI 369
Qy 350 LSTVGSYDWAGVFLYTSKEKSTFINMTRVDSDMN-DAVLGYAALILNRRVQSVILGAP 408
Db 370 VGSVGSNDWRGALYEVTG--SGSDFKETEIIDPAVNKDSYMGYSTVLGMRGVSLFLPSGAP 428
Qy 409 RYQHIGLVAMFRQNTGMESNANVRGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYTE 468
Db 429 RAHGTGLVTLFTKNQNTWTVTSNINEQIGSYFGASLSLLDSDSDSDFLLVGLAPLPYQ 488
Qy 469 QTRGQSVCPPLPGQRAKWCDAVLYGEQ-----GQPMGRFGAALTVLGDVNGKLTIV 523
Db 489 SQ-----PRTEGLYVYLSSEQYFKTLNVQSSTGTRFAASVASLKDNGDGLSDV 540
Qy 524 AIGAPGEDNRGAVYLFHGTSGSGISPSHS-ORIAAGSKLSPLRLQVFGQSLSGQDLTMDG 582
Db 541 AVGAPLE--NEGVIYILGDRTHGINPLTPORI SVQSVLPGLOQFGVSLTGQMDMNDN 598
Qy 583 LVDLTVAGAGHVLLRSQPLVKALMEENPREVARNPECDQVVKGEAGEVRVCLAV 642
Db 599 LTDIVIGAQGGIVLLKARFVMSVAQLSFSPKISILNYPECFGS--NAFNALNLTSCFTV 656
Qy 643 QKSTRDLREGQIQSV--VYDLDLDSGRPHRAVFNETKNSTR--RQTVLGLTQTCET 698
Db 657 TERTSS---TGSLEKLVNLVNLVDVVRGMRGTFDQSSVSSRTLQQSVLLDSGSSCFN 713
Qy 699 LKQLPNCIEDPVSPIVLRLNFSLVGTPLSAGNLRPVLAEDAQRILFTALPFPEKMGND 758
Db 714 PSIFMLRCVADTVSPKIRMFNS--QTEMLS-GNSVAVLDVHSRTEENVEVFPQRCSN 770
Qy 759 NTCQDLSITFESFMSLDCLVCGPREENVTVTVRNDGDSYRTQVTFPPPLDLSYRKVST 818
Db 771 NSCVADLKUNFSTN-NTLVENQAHFTVQVSLANPGDDSYNTSVLHVPEGISLSKPA 829
Qy 819 LQNRQSRSRWLACHSASSTEVSGALKSTCSINHPIPENSEVTFNITFDV---DSKAS 875
Db 830 IKPSRTR-----SSCGDRDSGATNRTTCSIDLFPVRSAGTTTQPLGTFRWKMDNDS 881
Qy 876 LGKULLLKXAVTSENNMPTNKTETFOELPKVYAVYVMTVSHGV-STKYLNFTASENTR 934
Db 882 NRMEMITANSNNGNM---SDTEVRSVFGVQAVDLAISLVAEDSVTFNFSLEDGRGPK 938
Qy 935 VMQHOYQVSNLQORSPLSLVFLVPR-----LNTPTFMSHEVHHSFSSYHQIIMCLIN 963
Db 939 ELNITYKVENSGKDLFVSVTLTLPCQPHVILTPTFMSHEVHHSFSSYHQIIMCLIN 998
Qy 964 QTVIWDPRQVTESEMLSSCTCHKEKRLPSSHDFLAKRPVNCSTAVCQRIQ-----C 1017
Db 999 KILLFPFSLAVQVQVRSTGMSLREVC-SQP--DLNKSVAHFNLTDARLQNVKEYS 1055
Qy 1018 DIPFFGICQEB--FNATLKNLSFDVYIKTSENHLLIVSTAELFNDVSFTLLPGQAFVR 1075

```

```
Db 1056 KYSFYEFKRDNVFSISAEINYNLSYNQTS-----SELKYNPH-----R 1094
Qy 1076 SQTETKVEFEVFNPEPLIV-GSSVCGLLLLALITLAALYKLGFFKQYKD 1124
Db 1095 SOTEKVKV-FVVPSPSLMLIVCTGAUGGVFFLLIILFLKCGCFKRNRPD 1143

RESULT 10
O96HB1 PRELIMINARY; PRT; 1086 AA.
AC Q96HB1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008777; AA08777.1; -.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SMC0191; Int_alpha; 5.
DR SMART; SMC0327; VWA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Hypothetical protein.
SQ SEQUENCE 1086 AA; 119223 MW; P6F2546B8C632F9 CRC64;

Query Match 23.0%; Score 1350.5; DB 4; Length 1086;
Best Local Similarity 32.1%; Pred. No 2.4e-30;
Matches 375; Conservative 187; Mismatches 432; Indels 175; Gaps 37;

Qy 1 FNLDTENAMTFQ--ENARFGQSQVWLOGSRVVGAPQEIIVANQSGSYQCDYSTGSC 58
Db 26 YNLDRVAGRSFPFRAGRHFGRVQLV-GNGVIVGAPGE---GNSTGSLYQCQSGTGCHL 81
Qy 59 PLELOVPEAVNMSLGLSLAATSPQLLACGFTVCTCSENTYVKGCLFGLGSLNRQOP 118
Db 82 EVTLR-GSNYSKLGMTLA--TDP----- 103
Qy 119 QKPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKEWSTVMEQLKSKTLFSLMQYSE 178
Db 104 -----TDGS----- 117
Qy 179 EFRHITFKFQNPNPRLIKPIITOLGRTHATGLRKVVRLNITNGARKAKILF 238
Db 118 SYKTEFDSDYVRKDPDALLKGVKMLLTNTFGAINVYATVFRFELGARPDAKVL 177
Qy 239 LLTDGKFGDPLGVEDVIPLEDRGVIRYVIGVDAPRSEKSRQELNLTVAKPPDRHVFQ 298
Db 178 IITDGE--ATDSGNIDAADK-----IIRVIIGIKHFQKESQETLHKFASKPASEFKI 230
Qy 299 INNFEALKTIONLRKIFAIKGTQTCGSSSPHEMSQBGFSAAITNSGPLSLTSGSYDW 358
Db 231 LDTFEKLKDLFTBLOKKIYVIEGTSQDLTFSNWELSSSGISADLSRGAHVAVGAKDM 290
Qy 359 AGGVF-LYTSKSKSTFTNTRVSDMNDVAVGLYAAA-IILRNVSQSLVLGAPRYQHIGLV 416
Db 291 AGGFLDLKADLQDDTTIGNEPLEPTEVRAGYLGITVTLWLSRQKTSLLASGAPRYQENGRV 350
```

```
Qy 417 AMFR--QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGG 474
Db 351 LLFREPQGGHWSQVTHGTQIGSYFGGELCGVDVDODGETELLIGAPFYGEQGR 410
Qy 475 VSVCLPRGQARMQCDAY--LYGEQGPWRGFAALTVLGDVNGDKLTDVAIGAPGED 532
Db 411 VFIV-----QRRQLGFEVSELOQDGPYPLGRFGEAITALTIDINGDGLVDVAVGAPLEB- 464
Qy 533 NRGAVYLPFGHTSGSGISPSHSORJAGSKLSPLQVFGOSLSGSGQDLTWDGLDVLTVGAQG 592
Db 465 -QGAVYIFNGRHG-GLSQPSQRLEGTOVLGSIQWFGRSIHGVRDLEGGDLADVAVGAE 522
Qy 593 HVLLRSQPLVRKAIMEFNPFRVARNVPECNDQV--KGKAGEVVRVCLHVQKSTRDLRL 651
Db 523 QMIVLSRPVDMVTLMSFSPAIPVHEVECSYSTSNKMKGVNITICFQI-KSLIPQF- 580
Qy 652 EGQIQSVVYDIALDSGRPHSAVNETKSTRQTVLGLTQCETLKLQPCNCEIDPV 711
Db 581 QGRLVANLTYTLQLDGHRTRRGPLPFGGRHLRNIAVT--TSMCTDTSFFHFPVCVQDLI 639
Qy 712 SPIVLRNLSL---VGTPLSAFEN-----LRPLAEDAQLFTALFPFKKNCGNNDICQ 762
Db 640 SPINVSINFLSWEBSGTTRDQORAGKDIPILRPSLSHSETWEI-----PFEXKCGDKKE 694
Qy 763 DLSITPFSMLDCLVCGPREFNTVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQVQ 822
Db 695 ANLRVSFSPARSRALRLTAFASLSVELSLNLEEDAYVQLDLHPFPGLSFKVEML--- 751
Qy 823 RSQSRWLACES--ASSTEVSGALKSTCSINHPIFFENSEVTFNITFDVDSKASLGNKL 880
Db 752 KPHSQIPVSCHELPEESLLSRAL---SCNVSSPIFKAGHSVALQMMNTLVNSWGDV 808
Qy 881 LKANVTSENN---MPTNKTEFOLELPVYAVVMVTSHGVSSTKYLNFTASENTSRV 936
Db 809 ELHANVTGNEDSDLLDENSATTI---IPILYPINILIQQDOBDSTLYVSTPKPKIHQV 865
Qy 937 QHQYQV---SNLQGRSLP-LSLVFLAVPVLQVTLWDRPQVTFSENLASTGHTK--ERLP 990
Db 866 KMYQVRIQPSIHDNIPTLEAVGVQPPSGPTTHQWSVQMEPPV--PCHYEDLERLP 923
Qy 991 SHSD--FLAELRKAPVNVNCSIAVCQRIQCDIPFGIOEENATLKNLSFDWYIKTSHNH 1048
Db 924 DAABECLPGLAFRCFVW-----FRQELVQVIGTGLVGEIAS-SM 964
Qy 1049 LLIVSTAILFNDVSTVFTLLPGCGAFVRQVETKVPPEVFPNPLIVGSSVGLLLALI 1108
Db 965 FSLCSLSISFNSKHFHLYGSNASL-AQVWVKVDVYVEKOKMLYLVLSGIGGLLLLLLI 1023
Qy 1109 TAALYKLGFPKQYKDMMSSEG-GPFGAP 1136
Db 1024 FIVLYKVGFFKENLKEKMEAGRGVNGIP 1052

RESULT 11
Q98TF0 PRELIMINARY; PRT; 1187 AA.
AC Q98TF0;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE CD11-2.
OS Cyprinus carpio (Common carp).
GN Cyprinus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Fujiki K., Nakao M.;
RT "Molecular cloning of a leukocyte integrin from the common carp.";
```



```
QY 174 MOYSEBFRTHFTKPEFONNPNRSLIKPTITOLLGRTHATGLRKVVRELFTNITNGARKNA 233
DB 4 VQSTYFRTEFTFLDIROKDPDALLAGYKXKELLTNTFGALNYVAKVFRDLPDARPA 63
QY 234 FKILFLLTGDKRPGDPLGYEDVPELDRGVTRYVIGVGDAPRSKSKROELMTVASKPPR 293
DB 64 TKVLIITIDGK---PPTNTTLARPKTSS---RSLIGIKNFKTKTESQALHQFASKPVE 116
QY 294 DHVFOINNFALTIQNLREKIFALEGTQTGSSSFHEHMQEGFSAAITSGPLLSV 353
DB 117 EFVKILDTTEKUKDLTEFKLIVJEGTSKQDLTSFNMELSSGISADLSGHHGVGVAV 176
QY 354 GSYDWAGGVF-LYTSKEKSTFFINMTVSDMDNDAYLYAAA-IILNRVQSLVIGAPRYQ 411
DB 177 GAKDWAGGFLDLKADLKSSTFFVGNBPLTVESRAGYLVTVTRLPSTGTMILLATGAPKYQ 236
QY 412 HGLVAMFRQ--NTGWESNANVKTOIGAYCASCISVDVDSNGSTDLVLCABHYEQ 469
DB 237 HNGRVLLFQPKRGPGPWSQIETDGIQISYFGELCGVDVDRDGGTELLLTAAPLYGE 296
QY 470 TRGGQVSVCLPRGQARWQCDVLYGEOQPMGRFGAALTIVLGDVNGDKLTDVAIGAFG 529
DB 297 QRGGRVPIY---QKIQLEQMSELQGETGYPLGRFGAIAALTIDINGDELTDVAVGAPL 353
QY 530 EEDNRGAVLPHGTSGSGISPSHSQIAGSKLSPRLQYFGOSLGGQDLTMDGLVLTG 589
DB 354 EE--QGAIVIFNGQOG-GLSPRESRIEQTGMPFSGIQWFGSTHGVKDLGGDLADVAVG 410
QY 590 AQGHVLLRSOPVLRVKAIMEFNPREVARVFNCDQVVKGEAG-EVRCVLHVQK--ST 646
DB 411 AEGQVIVLSRPVVDIITSVSFPALIPVHEVECSYSTSNQKKEGVNLVCTQVKSLIST 470
QY 647 RDLREGQIQSVVYTDALDSGPHRAVFNFTKSTRTQTVLGTQTCETIKLQLPNC 706
DB 471 ----FQGLVANYLTITLQDGHTRSRGLPFGGKHLIGNTAVTPV-KSCFVFWHFPIIC 525
QY 707 IEDPVSPVILRNPSL---VGTPLS--AFGNLPEVLAEDAQLFTALPFEKNCGNMNIC 761
DB 526 IQDLISPINVLSYSLWEEBGTDRPALDRDIPILKPSPHLETKEIIFPKNCGBDXNC 585
QY 762 QDDLSTFSPMSLDCLVWVGPRFNVTVTVRNDGDSYRQTQVTFPPDLDSYRKYSTLQ 821
DB 586 EADLKAFSDMRSKILRLTSPASLSVRLTLRNTAEDAYVWQVTLSPFQGLSPRKVBIL-- 643
QY 822 QRSQRWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITPOVDSKASLGNKL 881
DB 644 -KPHSHVPVGCBELPEAVVHS-RAUSCNVSSPIFGEDSMVDIQVMENTLQKSGWGDPIE 701
QY 882 LKANVTS-----ENNMPRTNKTBFLQLPVPKYAVYVTVTSHGVSTKYLNFTASENTS 933
DB 702 LOANVCNNESSILENSATTS-----IPWYPIVNLTKDQENSTLYISFPKSPRI 754
QY 934 RVMOHQYQV----SNLQORSLPLSLVPLVPLVRLNQTVI----WD---RPQVTFE-ENLST 982
DB 755 HHVKHIYQVRIQPSVNDMP-PLEALVVRVPRVHSEGLITHKWSIQMBEPVNCSPRLNESP 813
QY 983 CHTKERLPSHSDFLAELRKAPVNCSTAVCQRIQCDIPFGIOEBFNATLKGKLSFDWYI 1042
DB 814 SDENAE-----SCSPGT--EPRCPIDF--RQELIVQVNGAVELRGYI 850
QY 1043 KTSNHLIIVSTABILNDVFTLLPQAGFVRSQTEKTEKVEPEVNPPLIVGSSVGGL 1102
DB 851 KAS-SMLSLCSLAISSNKKHFLHGRNASM-AQVVMKYDLVYKEMLYLYVLSIGGL 908
QY 1103 LLLALITAAALYKLGFFKR 1120
DB 909 LLLFLIFIALYKVGFFKR 926
PRELIMINARY; PRT; 1167 AA.
O88340
RESULT 13
O88340
ID O88340
```

Db 489 -BEDAFVRRLEGQMGSGYFGSVLCPVDIDNDGTDTDFLLVAAPFYHIRGREGRVYVQVPE 547
Qy 483 GQARWOCDAVLVGEQGWGRGAALTIVLDVNGDKLTDVAIGAP-----GEEDNRGA 536
Db 548 -QDASFSLAHTLSGHPCLNRSRFGFMAAAGVDINQDKFTDVAIGAPLEGFGAGDGASYS 606
Qy 537 VYLFHGTSGSGISPSHSQIRAGSKLSPRLQYFQSLSGGQDLTMDGLVLTIVGAQGHVLL 596
Db 607 VYIYNGHSG-GLYDSFQQIRASSVASGLHYFGMSVSGGLDFNGDGLADITVGRSDSAV 565
Qy 597 LRSQPLRVKAINMFNPVRVARNVFCNDQVVKGEAGEVRVCLHVQKS---TRDRLREG 653
Db 666 LRSRPVVDLTVMSTFT?-----DALEMFVFGKV--DVNLCFEVDSSVASEPGLREM 715
Qy 654 QIQSVVYTDALDLSGRHSPAVNEKNSRTRTOVLGLTQTC-----696
Db 716 FLNFTVDV-----TKQRQLCEDSSGSCCLRWNGSGFLCEHFML 760
Qy 697 ETILQLQPNICIEDPVSPIVLRNFSIVGTPLSAGNLR-----PVLAEDAQRLFTALF-P 750
Db 761 STEEL-----CEDCFSNITIKVYE-----FQTSQGRRDYVNETL--DHYKEPSAIFQLP 809
Qy 751 FEKNCGNDNICQDLISITFSMSLDCLVGGPREFNVTVTVNDGEDSYRTQVTFPPFLD 810
Db 810 YESDCKNKVFCIAIQLITN-ISQQLVGVTVKEVTMNSLTSNGSDSYTNMAMLYPRN 868
Qy 811 LSYRKVSTLQNSQSRWSRLACESASSTVSGALKSTSCSINHPIPFENSEVTFNITFDV 870
Db 869 LQFKKI-----QKVPSPVQCDPXPV--ASVLVWCKIGHPII-KSSVNVSVTMQL 918
Qy 871 DSVASLGNKLLKANTSENNMPTNKTBFQLELPVKYAVMVTVTSHGVSTKYLNFATSE 930
Db 919 EESVFPNRTADITVTISNKSLSARETR--SLQFRHAFIAYLSR--PSVMYMN--TSQ 971
Qy 931 NTSRVNMOHOVENLQORSUPISLVFLVPVRLNQTWDRPOVTFPSNLSST-----CHT 985
Db 972 SPDSHKEFPNVGENLFGAVFQLOICVPIKQDP-----QIVRVKNLTKQDHTCTQ 1025
Qy 986 KERLPSSDFLABLRKAPVNVCSIAVCQRIQCDIPFGIQEIEBNATLKNLSPDWIKTS 1045
Db 1026 SQEPACGSDPVQHVHWSHVVCAI-----TSNKENVTVAEISVG 1065
Qy 1046 HNHLIIVSTA-----RILPNDVSTFLPGGAFVRSQETKVPEP-----EVENPLPLIV 1095
Db 1066 HTQLQIRDVSELPIGLSIFNKSLEYELNAB-----NHRKTIIVFLKEEETRSPLII 1119
Qy 1096 GSSVGGLLALITALYKLGFPKRYKDMWSE 1128
Db 1120 GSSIGGLLVLLWIIAILFKGFRKRYQQLMLE 1152

RESULT 14

O88341 PRELIMINARY; PRT; 1167 AA.
AC O88341;
DF 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Integrin alpha E2 (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brennan M., Rees D.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.,"
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020046; AAC23663.1; -
DR HSP; P11215; IBIHQ.

DR GO:0008305; C:integrin complex; IEA.
DR GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro: IPR000413; Intergrin_alpha.
DR Pfam: PF01839; FG-GAP; 3.
DR Pfam: PF00357; Integrin_A; 1.
DR Pfam: PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02334; VWEA; 1.
FT NON TER 1
SQ SEQUENCE 1167 AA; 128593 MW; E3FED7E18B5CBCEC CRC64;

Query Match 18.8%; Score 1109; DB 11; Length 1167;

Best Local Similarity 27.5%; Pred. No. 1.9e-72;

Matches 332; Conservative 206; Mismatches 42; Indels 244; Gaps 37;

Qy 47 LYQCDYSTGSCPTIRLQVPVBAVNMVSLGLSLAATTPSPQLLACGPTVH-----94

Db 67 LYQC-----ATSLIDEIGCCQVVEHLMPKGRYQGV 98

Qy 95 -----QTCSENTVVKGLCFELFQSNLRQOKPEALRG-----127

Db 99 LVRNHNGVLVCIQVQSRKPSRLNSBELTGACSLTLPNLDLQAGAYFSDLEGVLDLGASVNS 158

Qy 128 -----CPOED-----SDIAFLIDGSGSIIPHDFFRMKQWV 157

Db 159 GDIYSGKSGTGBETKSARRLRQAVBEDEEAGEATLAVLDGSGSIEPSDFOKAKOFI 218

Qy 158 STVMEQL--KSKTLFSLMOYSEBPRIHFTFKPQNNPNRSLIKPIITQLLQRTATGL 215

Db 219 STMENFYEKCFECFNALVQGVGIQTEFDLLDSDRINASLAKVQSIQVQKEVTTASAM 278

Qy 216 RKVPEELNITNGARKNAFKILFLLTDCGEKFGDPLGYEDVIPELDREGVIRVIGVDAF 275

Db 279 QVLDNITFIPGRSGRRKALKVWVLTGDI FRDPLNLTIVISSKMQGVVAFALGVGNAP 338

Qy 276 RSEKSRQELNIVASKPRDRHVQINNFEALXTI ONQLREKIPAEIGTGTGSSSSFEHMS 335

Db 339 ENNNTYRELKLIASDPKAAHTFKVTNYSALDGLLSKLQRIIHMBGT---VGDTLQYQLA 395

Qy 336 QEGPSAAITSNP-LLSTVGSVDNAGGVFLY-TSKEKSTFELNMTFVDSMDNA---YLGY 390

Db 396 QTFPSAQILDKGVLGTGVAFNWSGGALLTNTONGRFLNQT-AKEDFRAAQYSLGY 454

Qy 391 AAAIILNRVQSLVLAGAPYQHIGLVAMFRONTGMWESNANVKGTVQIGAYFASLCSVDV 450

Db 455 SVAALHKAHGSVYVAGAPRHLRGAVFELQKEDGETPMRRIEGBQMSYFGSVLCPVDI 514

Qy 451 DNGSTDLVLGAPHYETQGGQVSVCLPRGQARWQCDVLYGEOGQWGRFGAALT 510

Db 515 NMDGITDFLLVAAPFYHIRGREGRVYVYRV-HEQDAPFSLVYTSGLYSGPGLTSRSGFAMA 573

Qy 511 VLGDVNGDKLTDVAIGAP-----GEEDNRGAVLFGKTSGSGISPSHSQIRAGSKLSPR 564

Db 574 AVGDINQDKFTDVAIGAPLEGFGAGDGASYSVYIYNGHSG-GLHASPQQIRASSVALG 632

Qy 565 LOYFOQSISGGQDLTMDGLVLTIVGAQGHVLLRSQPLRVKAINMFNPVRVARNVFEKN 624

Db 633 LYVFGMSVSGGLDFSGDDLADITVGSQDVAVYLSRPVVDLTVMSTFTPDALP-----685

Qy 625 DQVVGKEAGEVRVCLHVQKS---TRDRLREGIQSVVYTDALDLSGRHSPAVNEKYN 681

Db 686 ---MAFKDNQVDELCKFVDSNAVPSEPGLRGMSINFTVDV-----TKQ 727

Qy 682 STRRQ-----TOVLGLTQTCETIKLQLPNCIEDPVSPIVLRNAPS 721

Db 728 KQRLQCADRSQCQCLMKWSGSSLCHEFGLISTEEL-----CEDDFCSNITIKVSYE 781

```
QY 722 LVGTPLSARGLRPLVLAEDAORLFTALP--PEKNCNDNICODDLSTFESMSLDCLVW 779
DB 782 F-QTSERNHNPIL--DHYKEPAIQLPTEKCKKVFIAISQILTAISOQD-LVW 837
QY 780 GGPREFNVTVVRNDEGDSYRTQVTFPPPLDLSYKRVSTLQQRORWRLACESASSTE 839
DB 838 GITKEVTWNISLNGSDSYMTNMLNVPRLNQFKI-----QKPLSPDIOCDPKEV- 890
QY 840 VSGALKSTSCSINHPIFPENSEVTNITFDVSKASLGKLLKXANVTSENNMPRTNKE 899
DB 891 --ASVLWCKLGHPL-KRSSVNSVTWQLEESIPNRTADITWISNKSLSARETH 947
QY 900 FOLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYVNSLQGRSLPISLVPLVP 959
DB 948 --SLOPRHAFVLSR--PSVMYMN--TSQSSDHKEPFFNVHGENHFGAVFQLQICVP 1000
QY 960 VPLNQTVHNDPQVTFSENLSST-CHTXKRLPSHSDFLAELKAPAVWCSIAVCQIOCD 1018
DB 1001 ETLRDL-----QIRVXHLTKTQART-----ECTSQEPTCGSPVQNV--- 1039
QY 1019 IPFFGIQEEFNATL-----KGNLSFDWYIKTSHNELLI--YSTAEIL-----PNDSVFTL 1066
DB 1040 -----BEWHSVICAITSKENVTVAASISMGHTKQLLRDISELQILGEISFNKSLYEG 1092
QY 1067 LPOGQAFVRSQETKVEPP-----EVPNPLIVGSSVGGHLLILALITAALYKLGFEKQY 1122
DB 1093 LNAE-----NHRKTIIVFLKEEKPHSLPLIGSSIGLLVIVIIATLFCGSPFKRY 1146
QY 1123 KDMWSE 1123
DB 1147 KQLNLE 1152

RESULT 15
Q7TQC3
ID Q7TQC3 PRELIMINARY; PRT; 1188 AA.
AC Q7TQC3;
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE All integrin.
GN ITGAL1.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RA Tiger C.-F., Popova S.N., Velling T., Ekman S., Forsberg E.,
RA Gullberg D.;
RT "allb1 integrin is important for mesenchymal cell function:
RT elimination of allb1 leads to dwarfism."
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RA Johansson M., Popova S.N.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AX124460; AAM62130.3;
XM Integrin.
SQ SEQUENCE 1188 AA; 133067 MW; 69B2BF20D75E4271 CRC64;

Query Match 17.7%; Score 1043.5; DB 11; Length 1188;
Best Local Similarity 28.1%; Pred. No. 1.4e-67;
Matches 349; Conservative 180; Mismatches 495; Indels 219; Gaps 46;

QY 1 FNEIDTENMTFO-ENARGQCSVWQ--LOGSR-VVVGAPQEIIVANQORSLYCCDYSGS 56
DB 23 FNMMDTRPRVIAQSAFFGYTQOHDIGSKWLWVGAPMETNGHOKTGDVTKCPVTQGN 82
QY 57 CEPRL-----QPVVEAVNMSLGLSAATSPQLIACGPTVHQTCSENTYVKGCLFLF 110
DB 83 CTKLNGLRVLNSVSEKDNMRGLSLATNPXNSFLACSLWSHCGSSYYTTGMSRV 142
```

```
QY 111 GSNLRQOPKQPEALRGCPQBDSDIAFLIDGSGSIIPHDFRMKEWST-----VMEQLKK- 166
DB 143 NSNFRPSKTVP-ALQRC-QTYMDIVVLDCGNSIYP-----WVEQHFILNILKCF 192
QY 167 ----SKTLPSLMQYSEFRTHFTKEFQNNPNPSLKIPIITOLLG-RHRTATGLKRVRE 221
DB 193 YIGPGQIQVIGVOGEDAVHBFHLNLYRSVDVVVEAAASHIEQRGTETRTAFGLEFAESE 252
QY 222 LPNITNGARKNAFKILPLLTGGEKFGDPLGYEDVPIBELDREGVIRYVIGV-----GDAPR 276
DB 253 AFQ--KGRKGAKKVMIVITDGESHDS-P-DLEKVIQSEKDNVTYAVAVLGYNNRRGIN 309
QY 277 SEKSRQELNVTVASPPRDHVQIINNFEALKTIONQLREKIPIAIEGTQTGSSSSFEHESQ 336
DB 310 PETFINEIKYIASPDGKFFNFVTDEALKDIVDALGDRIEFSLEGTNK-NETSGLEMSQ 368
QY 337 EGFSAATISNGLPSTVCSYDMAGGVFLYTSKEK-----STFINMTRVDSMDNDAIYGA 391
DB 369 TGFSSHVVDEGILLGAVDNGAVLKEYSAGKVIPIHRESYLAKFPELKHNVAYLGYT 428
QY 392 AAILLRNV-OSLVGAPRYOHIGLVAMF-RQNTGMESNANVGTQIGAYFGASLCSVD 449
DB 429 VTSVSSRQGRVYVAGAPRFNHTGKVILFSMHNRSLLTIHQALRGEQIGSYGSEITSD 488
QY 450 VDSNGSTDLVLIGAPHYEOTR-GSOYSVCLPRGQARWQCDVLYGEOQOPWGRFGAA 508
DB 489 VNDRVTDLVLGAPMFSEGRBRGKVTVNL---RQNRFYNGTLKDSHYSQONARFQSC 545
QY 509 LTVLGDVNGDKLTVAIGAPCEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSPQLQYF 568
DB 546 IASVQDLNODSVNDVWVGAPLEDSHRGAIYIFHGQ-TNILKXPMQRTASELAPLQHF 604
QY 569 QCSLSGGQDLTMDGLVDTVGAQGHVLLRSQPLRVKAIEMENPREVARNVF--ECNDQ 626
DB 605 GCSIHGQDLNEDGLVDLAVGALGNVAVLWAPVQVQINASLHFEPSKI--NIFHKDC--- 659
QY 627 VVKGEAGEVRVCL-----HVQSTRDLREGGQIQSVVYVLDALDGRPHSR 673
DB 660 ----KNGRDNATCLAAFLCPIPIFLAPHFOATVG-----IRYNATMDERYMPR 705
QY 674 AVFNETNS--TRQFOVLGLTQTCTELKLOPNCEBPDVSPVLELNFSLAVGTPLSAFGN 732
DB 706 AHLDEGGQFTNRAVLSSGQEHQCRINFHVL-DADYVXKPAVSFEVSELEDP----- 758
QY 733 LRPLVLAEDAORLFTALPPEKNCNDNICODDL----- 765
DB 759 NGPMLDNGMPTLRSVPFNGCNEDEHCVEDVLDAESDLTAMEYQCVLGRPAQDCS 818
QY 766 SITFSFMSLDCLVGCGPRFNVTVVRNDEGDSRTQVTFPPPLDLSYKRVSTLQNRSQ 825
DB 819 SYTLSFDTVTFIIBSTRERRVAEATLENRGENAYSAVLNISQSENLP--ASLIQKDDSD 876
QY 826 RSWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVSKASLGKLLKX 884
DB 877 NS--IECVNEER-----RLHKVCNVSYPPFRAKAKVAFRLDFFSKSVFLHLHLQHLGA 929
QY 885 NVTSENNMPRTNKEFQLELPVKYAVVMVTVSHGVSTKYLNFTASENTS-----RVM 936
DB 930 GSDSHEQDSTADDNTALLRFLHLYEADVLFTR---SSLSHFEVKAANSLSYDGIGPFF 986
QY 937 QHQVQVNLGQSLPISLVFL---VPV-----RL-----NOTVIW 968
DB 987 NCVFVKQNLG--FFPIHGVMKMTVPPIATROGNRLMLRDFDTQGNSTCNWNGNSTYR 1044
QY 969 DRPQVTFSENLSSTCHTKERLPSHSD---FLAELKAPVWNCSTIAVCQIOCDIPFFGIQ 1025
DB 1045 STPT---EEDLS---HAPQRNHSNDSVVSIIICNRLAP-----S 1077
QY 1026 BEFNATLGNLSPDWYIKTSHNLLIVSTAEI---LFNDSVPTLLPGQAFVRSQETKVK 1082
DB 1078 QSTSYLVGNLWLTSEKALKYRSLSKITVNAALQROFHSPFIP-----REEDPSRQ 1127
QY 1083 BPFEVFN-----PLPLIVGSSVGGLLLELALITAALYKLGFFK 1119
```

Db 1128 VTFEISKQEDWQVPIWIVGSTLGGLLLLALLVIALWKLGFVK 1170

Search completed: June 7, 2004, 17:16:21
Job time : 44.4889 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:54:39 ; Search time 49.4608 Seconds
(without alignments)
6495.175 Million cell updates/sec

Title: US-09-902-481B-4
Perfect score: 5884
Sequence: 1 FNLDTENAWTFQENARGFGQ.....FRRQYKDMVSEGGPPGABPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5852	99.5	1153	2	Aaw65090 Human Bet
2	5852	99.5	1153	3	Aab07360 Human Cbl
3	5852	99.5	1153	5	Aau80252 Human Int
4	5852	99.5	1153	5	Abg61469 Human Bet
5	5852	99.5	1153	5	Aaol4428 Integrin
6	5852	99.5	1153	7	Add25615 Binding d
7	5842	99.3	1153	2	Aar04136 Alpha sub
8	3450	58.6	1163	2	Aar07120 p150.95 a
9	3436	58.4	1163	3	Aaw65091 Human Bet
10	3436	58.4	1163	3	Aab07361 Human Cbl
11	3436	58.4	1163	5	Abg61470 Human Bet
12	3434	58.4	1163	6	Abu07406 Protein d
13	3401	57.8	1161	2	Aar78166 Human bet
14	3401	57.8	1161	2	Aaw23049 Human bet
15	3401	57.8	1161	2	Aaw57491 Human bet
16	3401	57.8	1161	2	Aaw65089 Human Bet
17	3401	57.8	1161	2	Aaw72825 Human alp
18	3401	57.8	1161	2	Aaw73342 Human alp
19	3401	57.8	1161	3	Aab07359 Human alp
20	3401	57.8	1161	5	Abg61468 Human Bet
21	3385.5	57.5	1161	2	Aaw23064 Human bet
22	3385.5	57.5	1161	2	Aaw65106 Human Bet
23	3385.5	57.5	1161	2	Aaw72837 Human alp
24	3385.5	57.5	1161	2	Aaw73343 Human alp
25	3385.5	57.5	1161	3	Aab07376 Human alp

ALIGNMENTS

RESULT 1

Aaw65090
ID Aaw65090 standard; protein; 1153 AA.

XX AC Aaw65090;

XX 28-SEP-1998 (first entry)

XX Human Beta-integrin CD11b subunit protein.

XX Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW type-I diabetes; atherosclerosis; multiple sclerosis; aschma;
KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;
KW rheumatoid arthritis.

XX Homo sapiens.

XX US5728533-A.

XX 17-MAR-1998.

XX 07-JUN-1995; 95US-00485618.

XX 23-DEC-1993; 93US-00173497.

XX 05-AUG-1994; 94US-00286889.

XX 21-DEC-1994; 94US-00362652.

XX (ICOS-) ICOS CORP.

XX Van Der Vieren M, Gallatin WM;

XX WPI; 1998-208565/18.

XX Screening assay for modulators of integrin binding - using immobilised or
labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.

XX Example 5; Fig 1A-D; 106pp; English.

XX This sequence represents a human beta-integrin CD11b subunit which is
used to describe a method for identifying compounds that modulate the
interaction of the beta-integrin alpha-d subunit with a binding partner
of alpha-d which involves contacting an alpha-d polypeptide with an alpha
-d binding partner, one of which is immobilised and the other of which is
labelled, in the presence of a test compound, and determining if the
compound affects binding between the alpha-d polypeptide and alpha-d
binding partner, where the alpha-d polypeptide is alpha-d or its fragment
comprising the cytoplasmic, transmembrane or extracellular domain of
alpha-d. Compounds that modulate alpha-d binding could be used to treat

26 3385.5 57.5 1161 5 ABG61485
27 3226.5 54.8 1161 2 AAR78169
28 3224.5 54.8 1161 2 AAW23062
29 3224.5 54.8 1161 2 AAW60004
30 3224.5 54.8 1161 2 AAW72824
31 3224.5 54.8 1161 3 AAB07374
32 3224.5 54.8 1161 5 ABG61483
33 3217.5 54.7 1161 2 AAW65104
34 3217.5 54.7 1161 2 AAW73345
35 3212 54.6 1161 2 AAW23061
36 3212 54.6 1161 2 AAW60003
37 3212 54.6 1161 2 AAW65103
38 3212 54.6 1161 2 AAW72836
39 3212 54.6 1161 2 AAW73347
40 3212 54.6 1161 3 AAB07373
41 3212 54.6 1161 5 ABG61482
42 3208 54.5 1161 2 AAR78168
43 3203.5 54.4 1151 2 AAW23059
44 3203.5 54.4 1151 2 AAW60001
45 3203.5 54.4 1151 2 AAW65101

ABG61485 Human Bet
Aar78169 Rat alpha
Aaw23062 Rat beta
Aaw60004 Rat alpha
Aaw72824 Rat alpha
Aab07374 Rat alpha
ABG61483 Rat Beta2
Aaw65104 Rat beta-
Aaw23061 Mouse bet
Aaw60003 Mouse alp
Aaw65103 Mouse bet
Aaw72836 Mouse alp
Aaw73347 Mouse alp
Aab07373 Mouse alp
ABG61482 Mouse bet
Aar78168 Mouse alp
Aaw23059 Rat beta
Aaw60001 Rat alpha
Aaw65101 Rat beta-

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
 CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
 CC and rheumatoid arthritis
 XX

SQ Sequence 1153 AA;

Query Match	99.5%;	Score 5852;	DB 2;	Length 1153;
Best Local Similarity	99.2%;	Pred. No. 0;		
Matches 1128;	Conservative 7;	Mismatches 2;	Indels 0;	Gaps 0;

Qy	1	FNLDTENAMTFOENARGFGQSVOLQSRVVVVGAPQEIIVAAANQSGSYQCDYSGSCPEI	60
Db	17	FNLDTENAMTFOENARGFGQSVOLQSRVVVVGAPQEIIVAAANQSGSYQCDYSGSCPEI	76
Qy	61	RLQPVFAVNMVSLGLSLAATTSPPOLLACGTVHTQTCSENTYVKGCLFGLFGLSNLRQPOK	120
Db	77	RLQPVFAVNMVSLGLSLAATTSPPOLLACGTVHTQTCSENTYVKGCLFGLFGLSNLRQPOK	136
Qy	121	FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFFRMKEWSTVMEQKLKSKTLFSLAQYSEEF	180
Db	137	FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFFRMKEWSTVMEQKLKSKTLFSLAQYSEEF	196
Qy	181	RIHFTFKFQNNPNRSLIKDITOLLGHTHTATGLRKVRELFNITGARKNAFKILPLL	240
Db	197	RIHFTFKFQNNPNRSLIKDITOLLGHTHTATGLRKVRELFNITGARKNAFKILPLL	256
Qy	241	TGGEKFGDPLGYEDVPELDRGVIRYVIGVDGDAFRSEKSRQELNTVASKPPRDHVFQIN	300
Db	257	TGGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFRSEKSRQELNTVASKPPRDHVFQIN	316
Qy	301	NFEALKTIONQRLREKIFAIBGTQGTSSSSPHEHMQSGFSAITNSGPELLSTVGSYDWAG	360
Db	317	NFEALKTIONQRLREKIFAIBGTQGTSSSSPHEHMQSGFSAITNSGPELLSTVGSYDWAG	376
Qy	361	GVFLYTSKEKSTFINMTVDSDMDNDAYLGVAALIIILNRVQSLVGLGAPRYOHIGLVAMFR	420
Db	377	GVFLYTSKEKSTFINMTVDSDMDNDAYLGVAALIIILNRVQSLVGLGAPRYOHIGLVAMFR	436
Qy	421	QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEYQTRGGQVSCPL	480
Db	437	QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEYQTRGGQVSCPL	496
Qy	481	PRGORARWQCDVLYGEGQCPWGFAGALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF	540
Db	497	PRGORARWQCDVLYGEGQCPWGFAGALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF	556
Qy	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGVAQGHVLLRSQ	600
Db	557	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGVAQGHVLLRSQ	616
Qy	601	PVLKVKATMEENPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRGLREGQIOSVVT	660
Db	617	PVLKVKATMEENPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRGLREGQIOSVVT	676
Qy	661	YDLALDGRPHSRVAFNETKNSRROTQVLGLTQTCETLKLQLENCIEDPVSPIVLRNF	720
Db	677	YDLALDGRPHSRVAFNETKNSRROTQVLGLTQTCETLKLQLENCIEDPVSPIVLRNF	736
Qy	721	SLVGTPLSAPGNLPPVLAEDQRLFTALFPPEKNGCNDNICQDLSLTFPSMISLDCLVWG	780
Db	737	SLVGTPLSAPGNLPPVLAEDQRLFTALFPPEKNGCNDNICQDLSLTFPSMISLDCLVWG	796
Qy	781	GPREFNVTVTVNRNDEGDSYRQVTFPPFDLSYRKVSTLQNRQSRQSRWLACESASSTEV	840
Db	797	GPREFNVTVTVNRNDEGDSYRQVTFPPFDLSYRKVSTLQNRQSRQSRWLACESASSTEV	856
Qy	841	SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKLLKANVTSENNMPKNTKTEF	900
Db	857	SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKLLKANVTSENNMPKNTKTEF	916
Qy	901	QLSEPVKYAVTAVTSHGVSTKYNFTASENTSRVMQHOVQVSNLQGRSPISLVLVPV	960
Db	917	QLSEPVKYAVTAVTSHGVSTKYNFTASENTSRVMQHOVQVSNLQGRSPISLVLVPV	976

RESULT 2

AAB07360 standard; protein; 1153 AA.

XX AAB07360;

DT 17-JAN-2001 (first entry)

XX Human CD11b protein sequence.

DE Human; macrophage infiltration inhibition; alpha d integrin;
 XX leukocyte integrin; Leu-CAM; leukointegrin; immune response;
 KW inflammation; leukocyte adhesion deficiency; IAD; Type I diabetes;
 KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
 KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
 KW rheumatoid arthritis; central nervous system injury; CD11b.

XX Homo sapiens.

XX WO200029446-A1.

XX 25-MAY-2000.

XX 16-NOV-1999; 99NO-US027139.

XX 16-NOV-1998; 98US-00193043.

XX 08-JUL-1999; 99US-00350259.

XX (ICOS-) ICOS CORP.

XX Gallatin MW, Van Der Vieren M;

XX WPI; 2000-387751/33.

XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
 macrophage infiltration and reduce inflammation at central nervous system
 injury sites.

XX Example 5; Fig 1; 270pp; English.

XX Integrins are a class of membrane-associated molecules that participate
 in cellular adhesion. Integrins are made up of an alpha subunit and a
 beta subunit. One class of human integrins are restricted to expression
 in white blood cells and have a common beta2 subunit: the leukocyte
 integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
 have an important role in immune and inflammatory responses. The present
 protein sequence is the human integrin alpha subunit CD11b. This sequence
 was used in an alignment to identify a novel beta2 integrin alpha
 subunit: alpha d (AAA60014 and AAB07359). The present sequence has
 approximately 60% identity to the protein sequence of alpha d. The
 alpha d gene and protein may be useful in therapy for diseases linked to
 aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple
 sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
 distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
 (IAD). In addition, anti-alpha d monoclonal antibodies may be used in the
 inhibition of macrophage infiltration at the site of a central nervous
 system injury. The monoclonal antibodies can also be used to detect and
 diagnose Crohn's disease

SO Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 3; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0

QY	1	FNLDTENAMTFOENARCFGQSVVOLQSGRVVVGAPQBIVAANORGSLYQCDYSTGSCPEI	60
DB	17	FNLDTENAMTFOENARCFGQSVVOLQSGRVVVGAPQBIVAANORGSLYQCDYSTGSCPEI	76
QY	61	RLQVPVEAVNMSLGLSLAAITTSPPQLAACPTVHQTCSENTYKVKLCFLFGSNLRQOPOK	120
DB	77	RLQVPVEAVNMSLGLSLAAITTSPPQLAACPTVHQTCSENTYKVKLCFLFGSNLRQOPOK	136
QY	121	FPEALRCPOEDSDIAFLIDGSGSIIIPHDPRMKEWYSTWEOIKKSKTILFSLMQYSEEF	180
DB	137	FPEALRCPOEDSDIAFLIDGSGSIIIPHDPRMKCBFYSTWEOIKKSKTILFSLMQYSEEF	196
QY	181	RIHFTKFEQONPNPRSLIKPIITOLLGRTHATGLRKVAVRELFINITNGARKNAFKILFLL	240
DB	197	RIHFTKFEQONPNPRSLIKPIITOLLGRTHATGIRKVVARELFNITNGARKNAFKILVVI	256
QY	241	TDEKFGDPIGYEDVIBELDREGVIRVIVGEDAPRSEKSRQBELNTVASPPRDRHVQIN	300
DB	257	TDEKFGDPIGYEDVIEADREGVIRVIVGEDAPRSEKSRQBELNTITASPPRDRHVQVN	316
QY	301	NFEALKTIQNLREKIFALIEGTQTGSSSSPEHEMSQEGFSAATISNGPLLISTVGSYDWAG	360
DB	317	NFEALKTIQNLREKIFALIEGTQTGSSSSPEHEMSQEGFSNAITSNGPLLISTVGSYDWAG	376
QY	361	GVFLYTSKEKSTFINMTVRVDSDMNDAYLGAAAIILNRVQSLVIGAPRYOHIGLVAMFR	420
DB	377	GVFLYTSKEKSTFINMTVRVDSDMNDAYLGAAAIILNRVQSLVIGAPRYOHIGLVAMFR	436
QY	421	QNTGMWESNANVKGTQIQAIFGASLCSVDVDSNGSTDVLIGAPHYBYOTRGGOVSVCP	480
DB	437	QNTGMWESNANVKGTQIQAIFGASLCSVDVDSNGSTDVLIGAPHYBYOTRGGOVSVCP	496
QY	481	PRGORARWQDAVLVYGQGPWGRFGAALTVLGVNVDGKLTUVAIGAPGEDNRGAYILF	540
DB	497	PRGORARWQDAVLVYGQGPWGRFGAALTVLGVNVDGKLTUVAIGAPGEDNRGAYILF	556
QY	541	HGTSGSGISPSHSORLIGSKLSPLROYFGQSLSGQDLTMDGLVDLTVGQGHVLLLRSQ	600
DB	557	HGTSGSGISPSHSORLIGSKLSPLROYFGQSLSGQDLTMDGLVDLTVGQGHVLLLRSQ	616
QY	601	PVLRVKAIMBFPNPREVARNVPECNDQVYVKGKEAGEVRYVCLHVQKSTRDLREGIOQSVVT	660
DB	617	PVLRVKAIMBFPNPREVARNVPECNDQVYVKGKEAGEVRYVCLHVQKSTRDLREGIOQSVVT	676
QY	661	YDLALDSGRPHSRVAVNETKNSTRQTVLGLTOTCETLKLQLEPNCIEDPVSPIVLRLNF	720
DB	677	YDLALDSGRPHSRVAVNETKNSTRQTVLGLTOTCETLKLQLEPNCIEDPVSPIVLRLNF	736
QY	721	SLVGTPLSAFGLNRPVLAEDAQRLLFTALPPEKXNCGNDNICQDDLSITFSFMSLDCIUVG	780
DB	737	SLVGTPLSAFGLNRPVLAEDAQRLLFTALPPEKXNCGNDNICQDDLSITFSFMSLDCIUVG	796
QY	781	GPREFNVTVVRNDGDSYRTQVTFPFLDLSYRKVSTLQNRQSQRWRKLACSSASSTEV	840
DB	797	GPREFNVTVVRNDGDSYRTQVTFPFLDLSYRKVSTLQNRQSQRWRKLACSSASSTEV	856
QY	841	SGALKSTSCSINHPIPPENSESVTFNITFDVDSKASLGNKULLKANVTSENMPRTNKTEF	900
DB	857	SGALKSTSCSINHPIPPENSESVTFNITFDVDSKASLGNKULLKANVTSENMPRTNKTEF	916
QY	901	QLELPPVKYAVVMVYTSHGVSFKYLNFTASENTSVRMQHOQVSNLQORSPLISVLPLVPV	960
DB	917	QLELPPVKYAVVMVYTSHGVSFKYLNFTASENTSVRMQHOQVSNLQORSPLISVLPLVPV	976
QY	961	RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVYVNCISIAVCORIQCDIP	1020
DB	977	RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVYVNCISIAVCORIQCDIP	1036

Qy	1021	FFGIQBEFNATLKGNISFDWYIKTSNNHLLIVSTAILFLNDSVFETLLPGQAFVRSOTET	1080
Db	1037	FFGIQBEFNATLKGNISFDWYIKTSNNHLLIVSTAILFLNDSVFETLLPGQAFVRSOTET	1096
Qy	1081	KVSEPFVNPPLPIVGSVSGLLILLALITAAALKLGFFKQYQKDMSEGGPPGAEBPQ	1137
Db	1097	KVSEPFVNPPLPIVGSVSGLLILLALITAAALKLGFFKQYQKDMSEGGPPGAEBPQ	1153

RESULT 3
AAU80252
ID AAU80252 standard; protein: 1153 AA.

AC BAU80252:

DT 15-III.-2002 (first entry)

XX DE Human integrin γ alpha-M subunit protein.

XX Integrin: antiinflammatory: immunosuppression: nephritis: dermatitis: XX

KW inflammatory disease; autoimmune disorder; chronic disease;
KW human immunodeficiency virus: HIV: myocardial infarction:

KW Sjorgen's syndrome; rheumatoid arthritis.

OS Homo sapiens.

FH	Key	Location/Qualifiers
mm	W-1	400
mm	W-2	400

FT /note= "Encoded by GGG CAG AGG"

PN WO200218583-A2.

PD 07-MAR-2002.

PF 31-AUG-2001; 2001WO-US027227.

PR 01-SEP-2000; 2000US-0229700P.

PA (BLOO-) CENT BLOOD RES INC.

PI Springer TA, Shimoaka M, Lu C;

DR WPI; 2002-382964/41.

XXXXXX

PT immunogen to produce antibodies specific to polypeptide, comprises a PT disulfide bond such that polypeptide is stabilized in a desired PT conformation.

PS Disclosure; Page 109-112; 112pp; English.

This invention relates to a modified integrin-I or integrin I-like domain polypeptide comprising at least one disulfide bond so that the domain is stabilised in a desired conformation. The polypeptide of the invention may have anti-inflammatory or immunosuppressive activities. The polypeptides of the invention have an open conformation and are useful as immunogens to produce antibodies that selectively bind to integrin I-domain; and for identifying a modulator of integrin activity, or of interaction of an integrin and a cognate ligand. The polypeptide of the invention, or antibodies (preferably anti-IFA-1 antibody) is useful for treating or preventing an integrin mediated disorder which is an inflammatory or autoimmune disorder in a subject and for inhibiting the binding of an integrin to a cognate ligand such as Crohn's disease, nephritis; human immunodeficiency virus (HIV), myocardial infarction, Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic composition comprising the peptide of the invention is useful for treating an integrin mediated disorder in a subject. The polypeptides and/or active or antigenic fragments are useful as reagents for diagnosis of integrin-mediated disorders. The present sequence represents the human integrin-1 alpha-M protein subunit used to generate the mutant polypeptides of the invention

XX	Sequence	1153	AA;
QY	Query Match	99.5%;	Score 5852; DB 5; Length 1153;
QY	Best Local Similarity	99.2%;	Pred. No. 0;
QY	Matches 1128; Conservative	7;	Mismatches 2; Indels 0; Gaps 0;
QY	1	FNLDTENAMTFQBNARGFGQSVVOLQSGRVVVGAPQBIIVAANORGSLYQCDYSTGSCPEI	60
DB	17	FNLDTENAMTFQBNARGFGQSVVOLQSGRVVVGAPQBIIVAANORGSLYQCDYSTGSCPEI	76
QY	61	RLQVPVEAVNMSLGLSLAANTSPOLLACGPTVHQTCSNTYVKGICFLFGSLNRQOPQK	120
DB	77	RLQVPVEAVNMSLGLSLAANTSPOLLACGPTVHQTCSNTYVKGICFLFGSLNRQOPQK	136
QY	121	FPBALRGCPQEDSDIAFLIDGSGSIIPHDPRRMKEWYSTMBOLEKSKTLFSLMQYSEEP	180
DB	137	FPBALRGCPQEDSDIAFLIDGSGSIIPHDPRRMKEFVSTMBOLEKSKTLFSLMQYSEEP	196
QY	181	RIHFTKPEQNNPNPSRLIPIQTOLLGRTHTAGLRKVRLELNI TNGAKNAFKILFLL	240
DB	197	RIHFTKPEQNNPNPSRLIPIQTOLLGRTHTAGLRKVRLELNI TNGAKNAFKILFLL	256
QY	241	TDGEKFGDPIGYEDVPELDRREGVIRVIVGWDAFRSEKSRQELNTVASKPRDRHVFQIN	300
DB	257	TDGEKFGDPIGYEDVPEADREBGRVIRVIVGWDAFRSEKSRQELNTVASKPRDRHVFQIN	316
QY	301	NPEALTKTIQNLREKIPAIISGTQTGSSSSSEHEMSQSGPSAAITSGNPILLSTVGSYDNAG	360
DB	317	NPEALTKTIQNLREKIPAIISGTQTGSSSSSEHEMSQSGPSAAITSGNPILLSTVGSYDNAG	376
QY	361	GVFLYTSKESKSTFINMTRVSDMNDAYLGAAGAAIILNRRVQSLVLAGAPRYOHIHLVAMFR	420
DB	377	GVFLYTSKESKSTFINMTRVSDMNDAYLGAAGAAIILNRRVQSLVLAGAPRYOHIHLVAMFR	436
QY	421	QNTGMMESNANVXGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQGVSVCP	480
DB	437	QNTGMMESNANVXGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQGVSVCP	496
QY	481	PRGORAPWQCDAVLYGEGQGPWGFAGALTVILGDVNDCKLTDVAIGAPGEDNRGAVYLF	540
DB	497	PRGORAPWQCDAVLYGEGQGPWGFAGALTVILGDVNDCKLTDVAIGAPGEDNRGAVYLF	556
QY	541	HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ	600
DB	557	HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ	616
QY	601	PVLKRVKALMEFNPREVARNVPEQNDQVYVKGKEAGEVRVCLHVQKSTDRERQIQSVWT	660
DB	617	PVLKRVKALMEFNPREVARNVPEQNDQVYVKGKEAGEVRVCLHVQKSTDRERQIQSVWT	676
QY	661	YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRINF	720
DB	677	YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRINF	736
QY	721	SLVGTPLSAPGNLAPVLAEDAQRLLFTALPFPEKNCGNNDI CODDLSITPSFMSLDCLVWG	780
DB	737	SLVGTPLSAPGNLAPVLAEDAQRLLFTALPFPEKNCGNNDI CODDLSITPSFMSLDCLVWG	796
QY	781	GPREFNVTVTRNDGDSYRTQVTFPPDLSTRKRVSTLQVRSQRORSWRLACSSASSTEV	840
DB	797	GPREFNVTVTRNDGDSYRTQVTFPPDLSTRKRVSTLQVRSQRORSWRLACSSASSTEV	856
QY	841	SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKXANVTSENMPRTNKTEF	900
DB	857	SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKXANVTSENMPRTNKTEF	916
QY	901	QLELPVKYAVVMVYTSHGVSXYLNFPTASENTSRVMQHOYQVSNLQORSIPISLVFLVPV	960
DB	917	QLELPVKYAVVMVYTSHGVSXYLNFPTASENTSRVMQHOYQVSNLQORSIPISLVFLVPV	976
QY	961	RLNQTIVWDRPQVTFSENLSSCTHTKRLPSHDFLAELRKAPVNVNCSIAVCORIQCDIP	1020

SQ		Sequence 1153 AA;	
Query Match		99.5%; Score 5852; DB 5; Length 1153;	
Best Local Similarity		99.2%; Pred. No. 0;	
Matches 1128; Conservative		7; Mismatches 2; Indels 0; Gaps 0;	
QY	1	FNLDTENAMTFOENARGPQSQVQVQGSRRVVGAPQEIIVAAANQORSLVQCDYSTGSCBPI	60
DB	17	FNLDTENAMTFOENARGPQSQVQVQGSRRVVGAPQEIIVAAANQORSLVQCDYSTGSCBPI	76
QY	61	RLQVPVEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVKGCLFGLSGNLRLQOQPK	120
DB	77	RLQVPVEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVKGCLFGLSGNLRLQOQPK	136
QY	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKBWVSTVMEQLKSKTFLSLMQYSEEF	180
DB	137	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKBWVSTVMEQLKSKTFLSLMQYSEEF	196
QY	181	RIHFTFKFQNNPNRSLKPIITQLLGRTHRTATGLRKVVRELFTNGARKNAFKILPIL	240
DB	197	RIHFTFKFQNNPNRSLKPIITQLLGRTHRTATGLRKVVRELFTNGARKNAFKILVVI	256
QY	241	TGEXFGDPLGYEDVPELDREGVIRYVTVIGVDADFSEKSRQELNTVASKPPDRDHVFQIN	300
DB	257	TGEXFGDPLGYEDVPELDREGVIRYVTVIGVDADFSEKSRQELNTVASKPPDRDHVFQIN	316
QY	301	NFEALXTIONLREKIPALEGCTQTGSSSFHEMSQEGFSAITNSGPLSTVGSYDWAG	360
DB	317	NFEALXTIONLREKIPALEGCTQTGSSSFHEMSQEGFSAITNSGPLSTVGSYDWAG	376
QY	361	GVFLYTSKEKSFPIINRTRVDSMDNDAYLGAAAILLRNRVQSLVGLGAPRYQHIGLVAFR	420
DB	377	GVFLYTSKEKSFPIINRTRVDSMDNDAYLGAAAILLRNRVQSLVGLGAPRYQHIGLVAFR	436
QY	421	QNTGMFESNANYKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTGQVSVCP	480
DB	437	QNTGMFESNANYKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTGQVSVCP	496
QY	481	PGQARWQCDVAVLGEQOPKRCGALTVLGDVNGDKLTVAICAPGEENRCAVILF	540
DB	497	PGQARWQCDVAVLGEQOPKRCGALTVLGDVNGDKLTVAICAPGEENRCAVILF	556
QY	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVLTGVAQGHVLLRSQ	600
DB	557	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVLTGVAQGHVLLRSQ	616
QY	601	PVLVRKAIIMEFPREVARNVFECDVQVKGKEAGEVRVCLHVOKSTRDLRRCQIOSVT	660
DB	617	PVLVRKAIIMEFPREVARNVFECDVQVKGKEAGEVRVCLHVOKSTRDLRRCQIOSVT	676
QY	661	YDLALDSGRPHGRAVFNKSTRQTVLGTQTCETLKLQLPNCIEDPVPSPVILRLNF	720
DB	677	YDLALDSGRPHGRAVFNKSTRQTVLGTQTCETLKLQLPNCIEDPVPSPVILRLNF	736
QY	721	SLVGTPLSAFGMLRPVLARDQRLFTALPPFKNGNDNICDDLSITPSFMSLCLVVG	780
DB	737	SLVGTPLSAFGMLRPVLARDQRLFTALPPFKNGNDNICDDLSITPSFMSLCLVVG	796
QY	781	GPREFNVTVVRNDGSDSYRTVTPFPPLDLSYRKVSTLQNRQSRQSNRLACESASSTEV	840
DB	797	GPREFNVTVVRNDGSDSYRTVTPFPPLDLSYRKVSTLQNRQSRQSNRLACESASSTEV	856
QY	841	SGALKSTCSGNIHPIPPENSEVTFTNTPDVDSKASLGNKLLKANVTSENMPRTKTEF	900
DB	857	SGALKSTCSGNIHPIPPENSEVTFTNTPDVDSKASLGNKLLKANVTSENMPRTKTEF	916
QY	901	QLELPVKYAVYVMTSHGVS TKYLNFTASENTRVMOHQYQVSNLQORSLPISLFLVLPV	960
DB	917	QLELPVKYAVYVMTSHGVS TKYLNFTASENTRVMOHQYQVSNLQORSLPISLFLVLPV	976
QY	961	RLNQTVINDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVAVNCISIAVQCRIQCDIP	1020
DB	977	RLNQTVINDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVAVNCISIAVQCRIQCDIP	1036
QY	1021	PFQIOEEFNATLKGNSLFDWYIKTSHNHLLIVSTAELFNDVSFTLLPGQCAFVRSQTE	1080
DB	1037	PFQIOEEFNATLKGNSLFDWYIKTSHNHLLIVSTAELFNDVSFTLLPGQCAFVRSQTE	1096
QY	1081	KVEPFEVNPFLVGVSSVGGILLALITAAALYKLGPFKROYKDMSEGGPPGAEPO	1137
DB	1097	KVEPFEVNPFLVGVSSVGGILLALITAAALYKLGPFKROYKDMSEGGPPGAEPO	1153
RESULT 5			
AA014428		standard; protein; 1153 AA.	
XX		AA014428;	
XX		03-MAY-2002 (first entry)	
DT		Integrin Mac-1 alpha subunit.	
DE		Mac-1; integrin alpha subunit; variant integrin inserted domain protein;	
XX		open conformation; integrin related inflammatory disorder;	
KW		integrin related immunological disorder; rheumatoid arthritis; ischaemia;	
KW		reperfusion; hypovolemic shock; infarction; cerebral shock;	
KW		viral infection; cancer; gene therapy; vaccine;	
KW		bioactive agent screening.	
XX		Unidentified.	
OS		WO200204521-A2.	
XX		17-JAN-2002.	
PD		09-JUL-2001; 2001WO-US021805.	
XX		07-JUL-2000; 2000US-0216600P.	
XX		(CALY) CALIFORNIA INST OF TECHNOLOGY.	
PA		(BLOO-) CENT BLOOD RES.	
XX		Springer T;	
XX		WPI; 2002-148167/19.	
DR		New integrin I domain protein having alteration in at least 2	
PT		noncontiguous regions and exits in an open conformation, useful for	
PT		treating, preventing or suppressing inflammatory or immunological	
PT		disorders.	
XX		Example 1; Fig 1F; 90pp; English.	
PS		The invention comprises structurally biased variant integrin inserted (I)	
XX		domain proteins, wherein the alterations to the protein occur in at least	
CC		two noncontiguous regions. Specifically the variant integrin I domain	
CC		proteins are structurally biased to exist in the open conformation.	
CC		thereby altering the binding ability of the protein. The invention also	
CC		comprises nucleic acids encoding the variant integrin I domain proteins.	
CC		The integrin I domain proteins and nucleic acids are useful for treating,	
CC		preventing or suppressing integrin related inflammatory and immunological	
CC		disorders (e.g. rheumatoid arthritis). The variant integrin I domain	
CC		proteins and nucleic acids can also be used for treating; ischaemia/	
CC		reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral	
CC		infection; and cancer. The variant integrin I domain nucleic acids and	
CC		proteins may be used in gene therapy, as vaccines and to screen for	
CC		bioactive agents. The present amino acid sequence represents the Mac-1	
CC		alpha subunit of integrin	
XX		Sequence 1153 AA;	
SQ		Query Match 99.5%; Score 5852; DB 5; Length 1153;	
Best Local Similarity		99.2%; Pred. No. 0;	
Matches 1128; Conservative		7; Mismatches 2; Indels 0; Gaps 0;	

QY 1 FNLDTENAMTFOENARGFGQSVVLQGSRRVVGAPQEI VAAQNRGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVLQGSRRVVGAPQEI VAAQNRGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLAGCPVTHQTCSENTYVKGICFLPGSNLRQOPK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLAGCPVTHQTCSENTYVKGICFLPGSNLRQOPK 136
QY 121 FPEALRGCCPOEDSDIAFLIDGSGSIIPHDPRKKEWVSTVMEQLEKSKTLPISLMQYSEBF 180
DB 137 FPEALRGCCPOEDSDIAFLIDGSGSIIPHDPRKKEWVSTVMEQLEKSKTLPISLMQYSEBF 196
QY 181 RIHFTKFEQNNPNPSLKIPIITQLLGRTHATGLRKVRELFNITNGARKNAKILFLL 240
DB 197 RIHFTKFEQNNPNPSLKIPIITQLLGRTHATGLRKVRELFNITNGARKNAKILFLL 256
QY 241 TDEKEGDPGLYEDVPEADREGVIRYVIGVGDAPFSEKSRQELNTVASKPRDHVQFN 300
DB 257 TDEKEGDPGLYEDVPEADREGVIRYVIGVGDAPFSEKSRQELNTVASKPRDHVQFN 316
QY 301 NFPAALKTIONOLEKIPAIETGOTGSSSPFHEHMSQEGFSAITNSGPIILSTVGSYDWA 360
DB 317 NFPAALKTIONOLEKIPAIETGOTGSSSPFHEHMSQEGFSAITNSGPIILSTVGSYDWA 376
QY 361 GVFLYTSKEKSTFINTRVDSNDNDAYLGYAAAIILNRVQSLVLCAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINTRVDSNDNDAYLGYAAAIILNRVQSLVLCAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVGTQIGYFAGSLCSVDVDSNGSDTLVLIGAPHYYEOTRGQSVCP 480
DB 437 QNTGMWESNANVGTQIGYFAGSLCSVDVDSNGSDTLVLIGAPHYYEOTRGQSVCP 496
QY 481 PRGORAWQCDVLYGEOGPWRCGALTVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 540
DB 497 PRGORAWQCDVLYGEOGPWRCGALTVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHESORAGSKLSPRLAYFGQSLGGQDLTMDGLDVLTVGAQGHVLLLR 600
DB 557 HGTSGSGISPSHESORAGSKLSPRLAYFGQSLGGQDLTMDGLDVLTVGAQGHVLLLR 616
QY 601 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVKSPDRRLRQIOSVVT 660
DB 617 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVKSPDRRLRQIOSVVT 676
QY 661 YDLALDSGRPHSAVFNKSTNRTOVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPHSAVFNKSTNRTOVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAFNGLRPVLAEDAQLFTALPFPKNGNDNICODDLSITFSKSLDCLVVG 780
DB 737 SLVGTPLSAFNGLRPVLAEDAQLFTALPFPKNGNDNICODDLSITFSKSLDCLVVG 796
QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACSSASTEV 840
DB 797 GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACSSASTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKANTVSENMPRTKTEF 900
DB 857 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKANTVSENMPRTKTEF 916
QY 901 QLELPVYAVYVTVSHGVSTKYNLFTASNTSRVMQHOVSNLQSRSLPISLVLPV 960
DB 917 QLELPVYAVYVTVSHGVSTKYNLFTASNTSRVMQHOVSNLQSRSLPISLVLPV 976
QY 961 RENQTVIWRDPQVTFSENLSSTCHTKERLPSSHDFLAELRKAQVWNCISAVCQRIQDIP 1020
DB 977 RENQTVIWRDPQVTFSENLSSTCHTKERLPSSHDFLAELRKAQVWNCISAVCQRIQDIP 1036
QY 1021 PFGQOEFNATLKNLSFDVWIKTSHHLLIVSTAEILLFNDVSFTLLPGQAFVRSOTET 1080
DB 1037 PFGQOEFNATLKNLSFDVWIKTSHHLLIVSTAEILLFNDVSFTLLPGQAFVRSOTET 1096
QY 1081 KVPEFVEPNPLTVGSSVGLLALLALITAAALYKLGFFKQYKDMMSGEGPQABPQ 1137

DB 1097 KVPEFVEPNPLTVGSSVGLLALLALITAAALYKLGFFKQYKDMMSGEGPQABPQ 1153
RESULT 6
ADD25615
ID ADD25615 standard; protein; 1153 AA.
XX AC ADD25615;
XX DT 15-JAN-2004 (first entry)
XX DE Binding domain-immunoglobulin fusion protein-associated protein #85.
XX KW Binding domain; immunoglobulin; fusion protein; cytostatic;
XX KW antiarthritic; immunosuppressive; antidiabetic; antithyroid;
XX KW neuroprotective; hinge region; immunoglobulin heavy chain;
XX KW CH2 constant region; CH3 constant region; IgG1;
XX KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
XX KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;
XX KW rheumatoid arthritis; myasthenia gravis; Grave's disease;
XX KW type 1 diabetes mellitus; multiple sclerosis; autoimmune disease.
XX OS Unidentified.
XX XX US2003118592-A1.
XX PN 26-JUN-2003.
XX PD 25-JUL-2002; 2002US-00207655.
XX PF 17-JAN-2001; 2001US-0367358P.
XX PR 17-JAN-2002; 2002US-00053530.
XX PR 03-JUN-2002; 2002US-0385691P.
XX XX (GENE-) GENE-CRAFT INC.
XX PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
XX DR WPI; 2003-801317/75.
XX XX New binding domain-immunoglobulin fusion protein, useful for treating a
XX PT subject having or suspected of having a malignant condition or a B-cell
XX PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.
XX PS Disclosure; SEQ ID NO 176; 157pp; English.
XX CC The invention relates to a binding domain-immunoglobulin fusion protein
XX CC comprising a binding domain polypeptide that is fused to an
XX CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain
XX CC CH2 constant region polypeptide that is fused to the hinge region
XX CC polypeptide, and an immunoglobulin heavy chain CH3 constant region
XX CC polypeptide that is fused to the CH2 constant region polypeptide. The
XX CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin
XX CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge
XX CC region polypeptide, derived from (a) having 3 or more cysteine residues;
XX CC where the mutated human IgG1 immunoglobulin hinge region polypeptide
XX CC contains 2 cysteine residues, where the first cysteine is not mutated; a
XX CC mutated human IgG1 immunoglobulin hinge region polypeptide, derived from
XX CC (a) having 3 or more cysteine residues, where the mutated human IgG1
XX CC immunoglobulin hinge region polypeptide contains no more than one
XX CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region
XX CC polypeptide, derived from (a) having 3 or more cysteine residues; where
XX CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains
XX CC no cysteine residues. The binding domain-immunoglobulin fusion protein is
XX CC capable of at least one immunological activity comprising antibody
XX CC binding domain polypeptide is capable of specifically binding to an
XX CC antigen. Also included are an isolated polynucleotide encoding the
XX CC binding domain-immunoglobulin fusion protein, a recombinant expression
XX CC construct comprising the polynucleotide operably linked to a promoter,
XX CC a host cell transfected or transfected with a recombinant expression
XX CC construct, producing the binding domain-immunoglobulin fusion protein, a

CC pharmaceutical composition comprising the binding domain-immunoglobulin
CC fusion protein or polynucleotide and a carrier and treating a subject
CC having or suspected of having a malignant condition or a B-cell disorder.
CC The binding domain-immunoglobulin fusion protein is useful for treating a
CC subject having or suspected of having a malignant condition or a B-cell
CC disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,
CC myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple
CC sclerosis or autoimmune disease. The present sequence is a binding domain
CC -immunoglobulin fusion protein-associated protein sequence. Note: The
CC sequence data for this patent formed part of the printed specification
CC and is also available in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not
CC identified the sequences in the printed specification by their SEQ ID
CC number therefore none of the sequences can be explicitly identified.
XX
SQ Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 7; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFENARGFGQSVVQLQGSVVVGAQOEIVAAHQRSGLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFENARGFGQSVVQLQGSVVVGAQOEIVAAHQRSGLYQCDYSTGSCBPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVXGCLFLFGSNLRQQPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVXGCLFLFGSNLRQQPQK 136
QY 121 PEARLGCPEQSDTAFLDGGSGIIPHDPRFKWGVSTVMEOLKSKTLPFLMOYSSEF 180
DB 137 PEARLGCPEQSDTAFLDGGSGIIPHDPRFKWGVSTVMEOLKSKTLPFLMOYSSEF 196
QY 181 RIHFTFKFQNNPNRSLKIPITQLLGRTHRTATGLRKKVRELNFNTNGARKNAKILFIL 240
DB 197 RIHFTFKFQNNPNRSLKIPITQLLGRTHRTATGLRKKVRELNFNTNGARKNAKILVVI 256
QY 241 TGEKFGDPLGYEDVITPELDRGVIRYVIGVADPRSEKSRQELNNTVASKPRDHVFQIN 300
DB 257 TGEKFGDPLGYEDVITPELDRGVIRYVIGVADPRSEKSRQELNNTVASKPRDHVFQIN 316
QY 301 NFEALKTIONQLREKIFALEGTQTGSSSFHEMSQEGFSAITNGPLLSVGVSYDWAG 360
DB 317 NFEALKTIONQLREKIFALEGTQTGSSSFHEMSQEGFSAITNGPLLSVGVSYDWAG 376
QY 361 GVFLYTSKSKSTFINNTRVDSOMNDAYLGAAAILLRNVQSLVIGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKSKSTFINNTRVDSOMNDAYLGAAAILLRNVQSLVIGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYETRGQVSVCPIL 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYETRGQVSVCPIL 496
QY 481 PRQQRARWQCDVLYGEOQGPWGRFGAALTVLGDVNGDKLTDAVIGAPCEEDNRGAVYLF 540
DB 497 PRQQRARWQCDVLYGEOQGPWGRFGAALTVLGDVNGDKLTDAVIGAPCEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGODLTMDGLVDLTVCAGQGVLLRSQ 600
DB 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGODLTMDGLVDLTVCAGQGVLLRSQ 616
QY 601 PVLRVKAIMFNPREVARNVFCNDQVKGKAGEVRVCLHVKQKSTRDLRGQIQSVVT 660
DB 617 PVLRVKAIMFNPREVARNVFCNDQVKGKAGEVRVCLHVKQKSTRDLRGQIQSVVT 676
QY 661 YDLALDSGRPHSRAVENETKNSTRQTVLGLTQTCETIKLQLPNCIEDPVSPVLRNLF 720
DB 677 YDLALDSGRPHSRAVENETKNSTRQTVLGLTQTCETIKLQLPNCIEDPVSPVLRNLF 736
QY 721 SLVGTPLSAPGNLRPVLAEDAQLFTALPFFPKNGCNDNICDDLSITFSFSLDCLVVG 780
DB 737 SLVGTPLSAPGNLRPVLAEDAQLFTALPFFPKNGCNDNICDDLSITFSFSLDCLVVG 796

RESULT 7
AAR04136

ID AAR04136 standard; protein; 1153 AA.
XX AAR04136;
AC AAR04136;
XX
DT 25-MAR-2003 (revised)
DT 07-SEP-1990 (first entry)
XX
XX Alpha subunit of Mac-1 leukocyte adhesion receptor.
KW Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
KW non-specific defence system; integrin gene superfamily.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FH Region 1..16
FT /label= signal peptide
FT Modified-site 86..88
FT /label= putative N-glycosylation site
FT Modified-site 240..242
FT /label= putative N-glycosylation site
FT Modified-site 391..393
FT /label= putative N-glycosylation site
FT Modified-site 469..471
FT /label= putative N-glycosylation site
FT Modified-site 693..695
FT /label= putative N-glycosylation site
FT Modified-site 697..699
FT /label= putative N-glycosylation site
FT Modified-site 735..737
FT /label= putative N-glycosylation site
FT Modified-site 802..804
FT /label= putative N-glycosylation site
FT Modified-site 881..883
FT /label= putative N-glycosylation site
FT Modified-site 901..903
FT /label= putative N-glycosylation site
FT Modified-site 912..914
FT /label= putative N-glycosylation site
FT Modified-site 941..943
FT /label= putative N-glycosylation site
FT Modified-site 947..949
FT /label= putative N-glycosylation site
FT Modified-site 979..981
FT /label= putative N-glycosylation site
FT Modified-site 994..996

QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 797 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFENSEVFNITFDVDSKASLGNKLLIKANVTSENNKMPSTNKTET 900
DB 857 SGALKSTSCSINHPIFENSEVFNITFDVDSKASLGNKLLIKANVTSENNKMPSTNKTET 916
QY 901 QLELPVKYAVVMVTVTSHGVSTKYLNFPTASENTRSVMOHQYQVSNLQGRSLPISLVFLVPV 960
DB 917 QLELPVKYAVVMVTVTSHGVSTKYLNFPTASENTRSVMOHQYQVSNLQGRSLPISLVFLVPV 976
QY 961 RLNQTIVDRPQVTFSENLSSTCHTKRSLPSHSDFLAELRKAPVWVNCISIAVCRIQCDIP 1020
DB 977 RLNQTIVDRPQVTFSENLSSTCHTKRSLPSHSDFLAELRKAPVWVNCISIAVCRIQCDIP 1036
QY 1021 PFGIOBEFNATLKGNSLSPFDWYIKTSHNHLIVSTABILLFNDVSFTLLPGQGFVRSQDET 1080
DB 1037 PFGIOBEFNATLKGNSLSPFDWYIKTSHNHLIVSTABILLFNDVSFTLLPGQGFVRSQDET 1096
QY 1081 KVPEPEVFNPLPLIVGSSVGGILLALITAAALYKLGFPFKQYKDMSEGPPGABPQ 1137
DB 1097 KVPEPEVFNPLPLIVGSSVGGILLALITAAALYKLGFPFKQYKDMSEGPPGABPQ 1153

FT Modified-site /label= putative N-glycosylation site
FT 1022..1024
FT /label= putative N-glycosylation site
FT 1045..1047
FT /label= putative N-glycosylation site
FT 1051..1053
FT /label= putative N-glycosylation site
FT 1076..1078
FT /label= putative N-glycosylation site
FT 1106..1134
FT /label= putative transmembrane region
XX
PN EP364690-A.
XX
XX 25-APR-1990.
XX
XX 17-AUG-1989; 89EP-00115159.
XX
XX 23-AUG-1988; 88US-00233533.
PR 09-MAR-1989; 89US-00321239.
XX
XX (DAND) DANA FARBER CANCER INST INC.
XX Springer TA, Corbi A;
XX WPI; 1990-125938/17.
DR N-PSDB; AAQ04043.
XX
XX New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating
XX inflammation and viral infections, and in diagnosis.
XX Disclosure; Page ?; -pp; English.
XX
XX Mac-1 alpha subunit is involved in the response to inflammation, i.e.
XX recognition of and migration to sites of inflammation. It also attaches
XX to cellular substrates as part of this function making it useful in
XX visualising endothelial tissue. Mac-1 is a member of the Integrin Gene
XX superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25
XX -MAR-2003 to correct PA field.)
XX
XX Sequence 1153 AA;
Query Match 99.3%; Score 5842; DB 2; Length 1153;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1127; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFOENARGFGQSVVLOGSRVVGAPQBIIVANORGSILYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVLOGSRVVGAPQBIIVANORGSILYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMISGLSLAATTPPQLACGPTVHQTCSNTYVKELCFGLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMISGLSLAATTPPQLACGPTVHQTCSNTYVKELCFGLFGSNLRQOPQK 136
Qy 121 PPEALRGCPQSDIAFLIDGSGIIPHDPRMKGVSTVMEQLKSKTFLSLMOYSEEF 180
Db 137 PPEALRGCPQSDIAFLIDGSGIIPHDPRMKGVSTVMEQLKSKTFLSLMOYSEEF 196
Qy 181 RIHFTFEFQNNPRSLIKPIITOLLGRTHATGLRVVRELFTNGARKNAPKILPL 240
Db 197 RIHFTFEFQNNPRSLIKPIITOLLGRTHATGLRVVRELFTNGARKNAPKILPL 256
Qy 241 TDGKFGDPLGYEDVIBELDREGVIRVVGDAFRSEKSRQELNTVASKPRDHVFPQIN 300
Db 257 TDGKFGDPLGYEDVIBELDREGVIRVVGDAFRSEKSRQELNTVASKPRDHVFPQIN 316
Qy 301 NFEALKTIONLRREKIPIAETGTQGTSSSPHEMSQGFSAATISNGPLLLSTVGSYDWAG 360
Db 317 NFEALKTIONLRREKIPIAETGTQGTSSSPHEMSQGFSAATISNGPLLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTVDSMDNDAYLVGAAAIIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVDSMDNDAYLVGAAAIIILNRVQSLVGLGAPRYOHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYYEQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYYEQTRGGQVSVCP 496
Qy 481 PRGQARWQCDVAVLYGEOGQPWGREGAALTIVLDVNGKLTVAIGAPGEEHNRGAVTLF 540
Db 497 PRGQARWQCDVAVLYGEOGQPWGREGAALTIVLDVNGKLTVAIGAPGEEHNRGAVTLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDILTMGCLVLTVAQAGHVLILASQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDILTMGCLVLTVAQAGHVLILASQ 616
Qy 601 PVLRVKAIKFNPREVARNVFECDNDQVYKGEAGVRVCLVHVKSTRDLREGQIOSVVT 660
Db 617 PVLRVKAIKFNPREVARNVFECDNDQVYKGEAGVRVCLVHVKSTRDLREGQIOSVVT 676
Qy 661 YDLALDSGRPHSRVFNETKSTRQTOVLGTCTETLKLQPCIEDPVPVILRLNP 720
Db 677 YDLALDSGRPHSRVFNETKSTRQTOVLGTCTETLKLQPCIEDPVPVILRLNP 736
Qy 721 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFPPEKNCNDNIQDDLSITPSPMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFPPEKNCNDNIQDDLSITPSPMSLDCLVVG 796
Qy 781 GPRENVTVTVRNDGEDSVRTQVTFPEPDLISYKRVSTLONORSORSNWLACESASSTEV 840
Db 797 GPRENVTVTVRNDGEDSVRTQVTFPEPDLISYKRVSTLONORSORSNWLACESASSTEV 856
Qy 841 SCALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKAVTSENMPRTNKTBF 900
Db 857 SCALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKAVTSENMPRTNKTBF 916
Qy 901 QLELPKVAIVYVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSLSPLVFLVVP 960
Db 917 QLELPKVAIVYVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSLSPLVFLVVP 976
Qy 961 RLNQTVINDRPOVTPSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
Db 977 RLNQTVINDRPOVTPSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1036
Qy 1021 PFGIOEFPNATLKGMLSPDWYIKTSHNHLIIIVSTAILFNDSDVFTLLPCGAFVRSQTET 1080
Db 1037 PFGIOEFPNATLKGMLSPDWYIKTSHNHLIIIVSTAILFNDSDVFTLLPCGAFVRSQTET 1096
Qy 1081 KVEPPRPVNPPLIIVGSSVGGILLALITAAALYKLGFFKRVQKDMWSEGGPPCAEPQ 1137
Db 1097 KVEPPRPVNPPLIIVGSSVGGILLALITAAALYKLGFFKRVQKDMWSEGGPPCAEPQ 1153
RESULT 8
AAR07120
ID AAR07120 standard; protein; 1163 AA.
XX
AC AAR07120;
XX
XX 25-MAR-2003 (revised)
DT 05-FEB-1991 (first entry)
XX
XX p150.95 alpha subunit encoded by clone lambdaX47.
XX
XX p150.95 leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia;
XX rhinovirus.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX 1..19 /label= signal peptide
XX 20..44 /label= N-terminus
XX 61..63 /label= glycosylation site
XX Modified-site

KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
KW lung inflammation; acute respiratory distress syndrome; CD11c subunit;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN US5728533-A.
XX
PD 17-MAR-1998.
XX
XX 07-JUN-1995; 95US-00485618.
XX
PF 23-DEC-1993; 93US-00173497.
XX
PR 05-AUG-1994; 94US-00286889.
XX
PR 21-DEC-1994; 94US-00362652.
XX
XX (ICOS-) ICOS CORP.
XX
XX Van Der Vieren M, Gallatin WM;
XX WPI; 1998-206565/18.
XX
XX Screening assay for modulators of integrin binding - using immobilised or
XX labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.
XX
XX Example 5; Fig 1A-D; 106pp; English.

CC This sequence represents a human beta-integrin CD11c subunit which is
CC used to describe a method for identifying compounds that modulate the
CC interaction of the beta-integrin alpha-d subunit with a binding partner
CC of alpha-d which involves contacting an alpha-d polypeptide with an alpha
CC -d binding partner, one of which is immobilised and the other of which is
CC labelled, in the presence of a test compound, and determining if the
CC compound affects binding between the alpha-d polypeptide and alpha-d
CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment
CC comprising the cytoplasmic, transmembrane or extracellular domain of
CC alpha-d. Compounds that modulate alpha-d binding could be used to treat
CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC and rheumatoid arthritis
XX
XX Sequence 1163 AA;

Query Match 58.4%; Score 3436; DB 2; Length 1163;
Best Local Similarity 60.6%; Pred. No. 1.2e-278;
Matches 684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;

QY 1 FNLDTENAMTQENARGFCQSVVQCGSRVVGAPQEIIVANQKSLQCDYSTGSCBPI 60
DB 1 FNLDTENAMTQENARGFCQSVVQCGSRVVGAPQEIIVANQKSLQCDYSTGSCBPI 79
QY 61 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCSNTYVKGLCFLFGSNLRQPOK 120
DB 80 GLQVPVEAVNMSLGLSLASTTSPSOLLACGPTVHCEGNNVLTGLCFLLGPT--QLTQR 137
QY 121 FPEALRGCEQSDIAFLIDGSGSIIPHDPRMKWVSTVMEQLKKSKTLFSLMOYSBEP 180
DB 138 LVSQRCEQPRQDVIPLIDGSGSSSRNFATMNFVRAVISQFQRPSTQFSLMOFSNKF 197
QY 181 RLHFTKEFQNNPNRSLKPTTOLLGRTHATGLKRVRELENTMGARKNAKIFLL 240
DB 198 QTHFTFEERPTSNPLSLASVHQLQGGTYTATATQNVVHRLPHASYGARDAIKILVI 257
QY 241 TPGKFGPDPLGYEDVNPIDLRGVIRYVIGVGDAFRSEKSRQELNVAASKPRDHVQIN 300
DB 258 TPGKKEGSDLYKDVIPADAAIIRYAIYVGLAFQNRNSWKELNDIAKSPSQEHIPKVE 317
QY 301 NPSALKTONLRKIFATBGTGCTGSSSFEHMSQEGFSAAITNGPILLSVGSYDWAQ 360
DB 318 DFDALNDIQNLRKIFALEGTETISSSSFELEMAQEGFSVFTPDGVLGAVGFTWSG 377
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVILGAPRYQHIGLVAMER 420

DB 378 GAFLYPPNMSPTFNNMSQENVDMDSDYLTGYSTELALWGVQSLVLAGPYQHIGKAVIFI 437
QY 421 QNTGHWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVILGAPHYEQTGGQSVQCP 480
DB 438 QVSRQWRMKAQVIGTQIGSYFGSLCSVDVDTGSDTLVILGAPHYEQTGGQSVQCP 497
QY 481 PRGQARMCDAVLYGEGQPGWRFGAALTIVLGVNVDKLTVDVAIGAFEGEDNRGAVYLF 540
DB 498 PRGWR-RWCDVLYGEGQPGWRFGAALTIVLGVNVDKLTVDVIGAPBENRGAVYLF 556
QY 541 HGTSGSISPSHSORIAGSKLSPLOYFGQSLSGQDLTMDGLVDLTVCAQGHVLLRSQ 600
DB 557 HGVLGPSISPSHSORIAGSKLSPLOYFGQSLSGQDLTMDGLVDLTVCAQGHVLLRSQ 616
QY 601 PVLAVKAIMBENPREVARNVEFECNDQVWKGAEAEVRVCLHVQSKSTRDLREGQTSVVT 660
DB 617 PVLAVGUSMQPIPAEIPRSAFECEQVVSQETLVQSNICLVYDKRSKNLGSRLDQSSVT 676
QY 661 YDLALDSGRPHSRVAVFNETKSTRRQVQLGTCTETLKQLPNCIEDPVSIVLRINP 720
DB 677 LDLALAPGRUSPRAIFQETKRSRVRVLGKAKHCENFNLLPSCVEDSVIPIRLNF 736
QY 721 SLVGTPLSAFCNLRPVLAEADQRLFTALFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 TLVGKPLAFNLRLPMLAALAQRYFTASLPEKNCGADHICQDNLGCSFSPGLKSLVG 796
QY 781 GPRFNVTVVRNDGDSYRTQVTFPPPLDLSYKVTQNLQNRQSRQSRWLACRSASSTEV 840
DB 797 SNLEINAEVWVWNDGDSYGTITFESHAPAGLSYRYVAEQKQGLRSLHLTC--CSAPVG 854
QY 841 SGALKSTSCSNHPIFFPENSVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
DB 855 SQGTWSTSCRNHLIFRGAQITFLATFDVSPKAVGLDRLLLIANVSENNIPRTSTIF 914
QY 901 QLELPVYAVVMYVTSHGVTYKILNPTAS-ENTSRVMOHQYQVGNLQGRSLPISLVLP 959
DB 915 QLELPVYAVYVYVSSHEQTKYLNFSSESEKSHVAMRYQVNVNGLQGRDLPVGINFWP 974
QY 960 VRLNQTVMIDRPOVTPFENLSSTCTKXERLPSHSDFLAELRKAPVNCVSIACVORICDI 1019
DB 975 VELNQBAMVMDVESHFQNPNSLRCSSEKIAPPASDFLAHQNPVLDOSIAGCLRFRCDV 1034
QY 1020 PFFGIGQEEFNATLKGSLSDFYIYKINTSHNHLIIVSTABILFNDVSFTLLPGQGFVRSQTE 1079
DB 1035 PSFSVQBELDFTLKGSLSGFWVRQILQKVSQVSVABIIIFDTSVYVSQLPGQEAFWAQT 1094
QY 1080 TRVFPPEVNPPLIVGSSVQGLLILALITLAALYKLGFFKQYKDMSE 1128
DB 1095 TVLEKYVHNPIPLIVGSSIGGLLLALITAVLYKVGFFKQYKEMKEE 1143
RESULT 10
AAB07361
ID AAB07361 standard; protein; 1163 AA.
XX
AC AAB07361;
XX
DT 17-JAN-2001 (first entry)
XX
DE Human CD11c protein sequence.
XX
KW Human; macrophage infiltration inhibition; alpha_d integrin;
KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;
KW inflammation; leukocyte adhesion deficiency; LAD; Type 1 diabetes;
KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
KW rheumatoid arthritis; central nervous system injury; CD11c.
XX
OS Homo sapiens.
XX
PN WO200029446-A1.
XX

PD 25-MAY-2000.
 XX 16-NOV-1999; 99WO-US027139.
 XX 16-NOV-1998; 98US-00193043.
 PR 08-JUL-1999; 99US-00350259.
 XX (ICOS-) ICOS CORP.
 PA Gallatin MW, Van Der Vieren M;
 XX MPI; 2000-387751/33.
 DR Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
 PT macrophage infiltration and reduce inflammation at central nervous system
 PT injury sites.
 PS Example 5; Fig 1; 270pp; English.
 XX Integrins are a class of membrane-associated molecules that participate
 CC in cellular adhesion. Integrins are made up of an alpha subunit and a
 CC beta subunit. One class of human integrins are restricted to expression
 CC in white blood cells and have a common beta2 subunit: the leukocyte
 CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
 CC have an important role in immune and inflammatory responses. The present
 CC protein sequence is the human integrin alpha subunit CD11c. This sequence
 CC was used in an alignment to identify a novel beta2 integrin alpha
 CC subunit: alpha d (AA60614 and AA607359). The present sequence has
 CC approximately 66% identity to the protein sequence of alpha d. The
 CC Alpha d gene and protein may be useful in therapy for diseases linked to
 CC aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple
 CC sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
 CC distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
 CC (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the
 CC inhibition of macrophage infiltration at the site of a central nervous
 CC system injury. The monoclonal antibodies can also be used to detect and
 CC diagnose Crohn's disease
 XX Sequence 1163 AA;
 SQ

Query Match 58.4%; Score 3436; DB 3; Length 1163;
 Best Local Similarity 60.6%; Pred. No. 1.2e-278;
 Matches 684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;

QY 1 FNLDZENAMTFQENARGFQSGVYVQGGSRVYVVGAPQEIIVAAHQRSGLYQCDYSTGSCBPI 60
 DB 20 FNLDTEELTAFRVDSAGFGDSVQVYVANSVYVVGAPQKIIAANQIGLYQCGYSTGACBPI 79
 QY 61 RLQVPVAVNNSLGLSLAATTPPOLLACGPTVHOTCSENTYVKGCLFLGSLNLRQOQK 120
 DB 80 GLQVPVAVNNSLGLSLASTTSPQLACGPTVHHCGRNMYLTGLCLLFGPT--QLTQR 137
 QY 121 FPEALRGCCQEBSDIAFLIDGSGSIIPHDPRMKWVSTVMEQLKKSKTLFSLMOYSEEF 180
 DB 138 LPVSRQECPRQDQIVFLIDGSGSISSRNFTATWAFVRAVISQFORPSTQFSLMQFSNKF 197
 QY 181 RHFTFKQNNPNRSLKPTOLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
 DB 198 QTHFTFEEFRRTSNPLSLASVHLQGGFTYTATQNVVHRLPHASYGARRDAIKILVI 257
 QY 241 TDGEKFGDPLGYEDVPELDREGVTRYVGVGDATSEKSRQELNATVASKPPRDVFOIN 300
 DB 258 TDGKEGSDLDNDVLPMDAAGIKYALGVGLAFQNRNSWELNDIASKPSQSEIFKVE 317
 QY 301 NFEALNTIQNLRKIFAIEGTQTSSSSFEHMSQEGFSAATISNGPLSTVGSYDAG 360
 DB 318 DEFALDKIQNLKIKIFAIEGTETITSSSFELEMAQEGFSAVFTPDGVLGAVGQFTWSG 377
 QY 361 GYFLYTSKESFTINRVDSDMDAVLYCYAAIITLRNVQSLAVGAPRYQHILVAMFR 420
 DB 378 GAFLYPPNMPSPFINNSQBNVMDRDSYGLSTELALWKGVSVLGAPRYQHIGRAVIFI 437
 QY 421 QNTGWESNANVKGITQIGAYFGASLCSDVDVDSNGSTDLVIGAPHYYEQTRGGQVSCVPL 480

Db 438 QVSRQMRKAEVIGTQIGSYFGASLCSDVDVDSNGSTDLVIGAPHYYEQTRGGQVSCVPL 497
 QY 481 PRGORARWOCDAVLYGEOGQFWRFGAALTVLGDVNGDKLTDVAIGAPGEDNREGAVYLF 540
 Db 498 PRGWR-RWVCDVLYGEOGHPGRFGAALTVLGDVNGDKLTDVVI GAPGEENREGAVYLF 556
 QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGAGCHVLLRSQ 600
 Db 557 HGVLPSPISPSHSQRIAGSKLSPRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLRTR 616
 QY 601 PVLRYKAIMEPNPREVARNVPECNQVVKGEAGEVRVCLHVOKSTDRRLREGOIQSVVT 660
 Db 617 PVLWVGSMQFIPAEIPRPAFCNEQVVSQTLVQSNICLYIDKRSKNLGRDLQSSVT 676
 QY 661 YDLALDSGRPHSRAVENETKSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
 Db 677 LDALAPGRLSRAIFQETKNSLSRVRLGLKAHCENFNLLPSCVEDSVIPIILRLNF 736
 QY 721 SLVCTPLSACNLRPVLAEQAORLPTALFPPEKNCNDNICODDLISITFSMSLDCLVVG 780
 Db 737 TLVGKPLAFNLRPLAALAQRYFTASLPFKNCAGADHICQDNLGIFSFFGLKSLIVG 796
 QY 781 GPREFNVTVTRNDGDSYRTQVTFPFLDLISYKVTQNLQNRQSRWRLACESASSTEV 840
 Db 797 SNLENAEAVMWNDGDSYGTITTFSPAGLSYRYVAEQKQGLRSLHLTC--CSAPVG 854
 QY 841 SCALKSTSCSNIHPIFENSEVTFNITFDVDSKASLGNKLLIKANVTSENNMPSTKTEF 900
 Db 855 SQGTWSTSCRINHLIFRGGAQITFLATFVSPKAVGLDRLLLIANVSENNIPRTSKTIF 914
 QY 901 QLELPVYAVVWVTVSHGVSTKYLNFAS-ENTSRVMQHOVQVSNLQORSIPLSLVLVP 959
 Db 915 QLELPVYAVVIVVSSHEQFTKYLNFSESEKSHVAMRYQVNVNLQORDLPVSINFVP 974
 QY 960 VRLNQTVIMDRPQVTFSENLSSTCHTVERLPDSDFLAELRKAVNCSIAVCORICDI 1019
 Db 975 VELNQEAVMMDVEVSHFPQNPSPSLRCSSEKIAAPPASDFLAHQNPVLDCSIAGCLFRCDV 1034
 QY 1020 PFGIGQEEFNATLKGNSLPDWYIKTSHNHLIVSTABILFNDSTVFTLLPGCGAFVRSQTE 1079
 Db 1035 PSFSVQELDFTLKGNLSFGWVRQILQKYSVVSVAIIIFDTSVYSLPQGEAFWEAQTI 1094
 QY 1080 TKVPPFEVFNPLPIVSGSVGLLLALITRALLYKLGFFKRYQKDMNSE 1128
 Db 1095 TVLEKYVHNPIPIVSGSSIGGLLLALITAVLYKGVGFKRYKEMMEE 1143
 RESULT 11
 ID ABG61470 standard; protein; 1163 AA.
 XX AC ABG61470;
 DT 27-AUG-2002 (first entry)
 XX Human Beta2 integrin alphaCD11c subunit.
 DE Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit; IAD;
 KW leukocyte adhesion deficiency; inflammatory response; diabetes;
 KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW immune complex alveolitis; leukaemia; ICAM-2; VCAM-1; anti-inflammatory;
 KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
 KW locomotor recovery; locomotor damage; locomotor impairment;
 KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
 OS Homo sapiens.
 XX WO200230980-A2.
 XX 18-APR-2002.
 PD 18-APR-2002.
 XX

PF 15-OCT-2001; 2001WO-US032059.
 XX 13-OCT-2000; 2000US-00688307.
 PR (ICOS-) ICOS CORP.
 XX Gallatin WM, Van Der Vieren M;
 XX WPI; 2002-463260/49.
 DR
 XX Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
 PT recovery, inhibiting locomotor damage, limiting locomotor impairment, or
 PT limiting autonomic and sensory dysfunction following spinal cord injury.
 XX Example 5; Page 194-198; 270pp; English.
 XX The invention relates to promoting locomotor recovery, inhibiting
 CC locomotor damage, limiting locomotor impairment, or limiting autonomic
 CC and sensory dysfunction following spinal cord injury by administering an
 CC anti-alpha d (beta2 integrin alpha2 subunit) monoclonal antibody to a
 CC spinal cord injury victim. The method also involves the use of a ligand
 CC selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule,
 CC vascular cell adhesion molecule). The method is useful for promoting
 CC locomotor recovery, inhibiting locomotor damage, limiting locomotor
 CC impairment, or limiting autonomic and sensory dysfunction following
 CC spinal cord injury. In particular, the spinal cord injury comprises
 CC compression of the spinal cord. The antibodies are also useful for
 CC reducing inflammation at the site of a central nervous system injury. The
 CC specification also details the identification of Beta2 integrin alphas
 CC cDNAs and proteins, for use in raising the antibodies. Beta2 integrins
 CC are implicated in diseases such as LAD (leukocyte adhesion deficiency,
 CC inflammatory response, diabetes, multiple sclerosis, arthritis, graft
 CC atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative
 CC colitis, immune complex alveolitis and leukaemia. The present sequence is
 CC a Beta2 integrin alpha subunit sequence included for comparison with the
 CC Beta2 integrin alphas protein sequences
 XX Sequence 1163 AA;
 SQ

Query Match 58.4%; Score 3436; DB 5; Length 1163;
 Best Local Similarity 60.6%; Pred. No. 1.2e-278;
 Matches 684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFQENARGCGSVVLOGSRVWVVCAPQIIVAAQGRSLVOCYDSTGSCBPI 60
 DB 20 FNLDTEBELTAFRVDSAGDSVVQYANSVWVVGAPQRIIAAQIGLYQCGYSTGACEFI 79
 QY 61 RLQVPVEAVNMISGLSLAATTSPOLLACGPTVHQCSENTYVYKGLCFLFGSNLRQOPQK 120
 DB 80 GLQVPEAVNMISGLSLAATTSPOLLACGPTVHQCSENTYVYKGLCFLFGSNLRQOPQK 137
 QY 121 PPEALRGCPQDSDIAFLDGSISIIPIHPRMKEWVSTVMEQLKSKTLPGLMYSSEEP 180
 DB 138 LPVSRQCPQEQDIYFLDGSISISRNPFATMMNFRAVISQFQSPSTQFSLMQFSNKP 197
 QY 181 RIHTPFKEFNNPNRSLKPIITQLGRTHATGLKRWVRELFINNGARKNAKILIFLL 240
 DB 198 QHTFTFEETPTNPSLLASVHQLQGYPTAITAIONVHRLPHASVGGARRDAIKILIVI 257
 QY 241 TDGKPGDPIGYEDVPELDRGVIRVYICVGDAFSEKSRQELNIVAKPRPDHVFQIN 300
 DB 258 TDGKPGDSDLDYKDVIPMDAAGIIRYAIGVGLAFQNRNSWKLNDIASKPSQEHIFKVE 317
 QY 301 NFEALKTIONQLREKFIATGCTGTGSSSFEHMSQGFSAATITNSGILLSTVGSVDWAG 360
 DB 318 DFDALXDIQNLKEKFIATGCTGTGSSSFEHMSQGFSAATITNSGILLSTVGSVDWAG 377
 QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLVGAAAIILNRVQSLVLGAPRYOHIGLVAMPR 420
 DB 378 GAFLYPPNMSPTFNNSQENVDSDYLGYSTELALWKGVSILVGLGAPRYOHIGKAVIFI 437
 QY 421 QNTGMVESNANVAGTQIGAFVFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQGVSCPL 480

DB 438 QVSRQWRKAEVICTQIGSYFGASLCSVDVDTGSDTDLVIGAPHYYEOTRGQGVSCPL 497
 QY 481 PRGORARWQCDVLYYGEQCPWGFAGALTVLGVNGDKLTDVAIGAPGEDNKGAVYLF 540
 DB 498 PRGWR-RWWDVLYYGEQCPWGFAGALTVLGVNGDKLTDVWIGAPGEDNKGAVYLF 556
 QY 541 HGTSGSISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
 DB 557 HGVLGPSISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
 QY 601 PVLRVKAIMENPREVARNVFECDNDVVVKGKEAEVRLVHVKSTRDRLEGOIOSVVT 660
 DB 617 FVLWVGSMQFIPAEIPRSAFECEQVWSEQLVQSNICLYIDIKRKNLGLSRDLQSSVT 676
 QY 661 YDLALDSGRPHSRVAFNETKNSRROTQVGLGTCTETLKLQLENCIEDPVSPIVRLNF 720
 DB 677 LDIALAPGLSPRAIFQETKNSRVRVVLGKHAHCENFNLLPSCVEDSVIPIILRLNF 736
 QY 721 SLVGTPLSAFONLPAEDAQRUFTALFPPEKNCNDNICODDLSTIFSPKSLDCLVVG 780
 DB 737 TLVGKPELLAFENLPLAALAQRYFTASLPPEKNCADHICQDNLGIFSFPGLKSLVVG 796
 QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYKRVSTLQNRQSRWRLACESASSTEV 840
 DB 797 SNLELNREVWVNDGEDSYGTITIFSHPAGLSYRYVAEGQKQGLRSLHLTC--CSAPVG 854
 QY 841 SCALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTET 900
 DB 855 SQGTWSTSCRINHLIFRGGAQITELATFDVSPKAVGLDRLILLIANVSENINPRTSKTIF 914
 QY 901 QLELPKYAVVMVTVSHGVSTKYLNFETAS-ENTSRVMQHQYQVSNLQORSIPISLVLPV 959
 DB 915 QLELPKYAVVMVTVSHGVSTKYLNFETAS-ENTSRVMQHQYQVSNLQORSIPISLVLPV 974
 QY 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFAELKAPVNVNCSIAVCORIQDI 1019
 DB 975 VELNQEAVMVDVEVSHFPQNSLRCSSEKIAPPASDFLAHQKPNFLDCLSIAGCLFRCDV 1034
 QY 1020 PFFGTOBEFNATLKNLSPDVAIVKTSNHLILIVSTABILENDVSFTLLPGQCAFVRSQTE 1079
 DB 1035 PSFVQBEELDTLKNLSFGVWRQRLQKKSWSVAEIIFTSVYSQLPGQAFWRAQTI 1094
 QY 1080 TKVPEFPEVNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKDMMSR 1128
 DB 1095 TVLEKYKHVNIPLIVGSSIGGLLLALITAVLYKVGFFKQYKEMMEE 1143
 RESULT 12
 ID ABU07406
 XX AC ABU07406;
 XX XX 28-JAN-2003 (first entry)
 XX DE Protein differentially regulated in prostate cancer #9.
 XX KW prostate cancer; gene expression; differential regulation;
 XX KW molecular marker; drug target; cancer detection; cancer diagnosis;
 XX KW cancer staging; cancer grading; cancer assessing; cancer monitoring.
 XX OS Homo sapiens.
 XX PN WO200281638-A2.
 XX PD 17-OCT-2002.
 XX PF 08-APR-2002; 2002WO-US010824.
 XX PR 06-APR-2001; 2001US-0281731P.
 XX PR 06-APR-2001; 2001US-0281732P.
 XX XX (ORIG-) ORIGENE TECHNOLOGIES INC.

FT Region /note= "extracellular domain"
 FT 150. .352
 FT /note= "this region is homologous to the insert common to
 FT ChIIa,b,c and may be a site for interaction with ICAM
 FT family proteins"
 FT 465. .474
 FT /note= "putative cation binding site"
 FT 518. .527
 FT /note= "putative cation binding site"
 FT 592. .600
 FT /note= "putative cation binding site"
 FT 1109. .1128
 FT /note= "transmembrane region"
 FT 1129. .1161
 FT /note= "cytoplasmic domain"
 FT XX
 PN W09517412-A1.
 XX 29-JUN-1995.
 XX 21-DEC-1994; 94WO-US014832.
 XX 23-DEC-1993; 93US-00173497.
 PR 05-AUG-1994; 94US-00286889.
 XX (ICOS-) ICOS CORP.
 PA Gallatin WM, Van Der Vieren M;
 PI WPI; 1995-240603/31.
 DR N-PSDB; AAQ91712.
 DR Alpha sub-unit polypeptide of human beta 2 integrin - used to identify
 PT potential antiinflammatory agents, for the treatment of graft
 PT arteriosclerosis, inflammatory bowel disease, asthma, etc.
 XX Claim 7; Page 82-87; 172pp; English.
 CC A probe based on a partial cDNA clone (given in AAQ91727) of canine alpha
 CC -TM1 was used to screen a human spleen cDNA library to identify clone
 CC 19A2 encoding the beta-1 integrin alpha-subunit, alpha-d. The cDNA was
 CC manipulated to allow expression of recombinant alpha-d subunit in COS and
 CC CHO cells
 XX SQ Sequence 1161 AA;
 Query Match 57.8%; Score 3401; DB 2; Length 1161;
 Best Local Similarity 59.3%; Pred. No. 1e-275;
 Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;
 QY 1 FNLDNENAMTQFNARFGQSVVQLGSRVVCAPQEIIVAAQORSLYOCYSTGSCERI 60
 DB 17 FNLDVREPTIFQDAGFGQSVVQFGSRLVVGAPLEVVAAQGTGRLYDCAATGMCQPI 76
 QY 61 RLOVPVEAVNMSGLSLAATTSPQACGPTVQCSTNTVYKGLCFLFGNLRQPOK 120
 DB 77 PLHIREAVNMSUGLTLASTNGSRLACGPTLHRCVGENSKSGCLLIGSEW-ELIOT 135
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHPRRNKEWVSTVMEQLKSKNTLSLMQVSEEF 180
 DB 136 VPDATPECPHQEMDIVFLIDGSGSIQDNDFNQMGFVQAVMGQFEGDTLFLALMOYSNLL 195
 QY 181 RHFTPEKQNNPNSLKIPIOLLGRTHATGLKRVVRELFINITNGARKNAKPILEL 240
 DB 196 KHFTPTQFTSQQSLSDVPIVOLKGLIFTATGILTVTQLPHKNGARKSAKILLIVI 255
 QY 241 TDGEKFGDPLGYEDVPIELDRGVIRYIVGVGDAFRSEKSRQELNVTASPRDHHVQIN 300
 DB 256 TDGQKYKDPLEYSVLPQAEKAGIIRYAGVGHAFQOPTARQELNTISSAPQDHHVFKVD 315
 QY 301 NFEALKTIQNLBEKIPAECTQTGSSSFEHEMSQEGSAITSGPLLSVTGSDVWAG 360
 DB 316 NFAALGSIQLOKEKYAVEGTQSRASSSSFOHEMSQEGFSTALTMQDLFLGAVGSPFWSG 375

QY 361 GVFLYTSKESKSTINMTRVSDMNDAYLGAAAILNRRVQSLVGLGAPRYCHIGLWAMER 420
 DB 376 GAFLYPPMPTFINMSQENVMDRDSYLGSTYALMWGVNQLVGLAPRQHTKAVIPT 435
 QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYHYEQTRGGQVSCPL 480
 DB 436 QVSRQWKKARVTCQIGSYFGASLCSVDVDSGSDTLILIGAPHYHYEQTRGGQVSCPL 495
 QY 481 PRGQARWQCDVLYGEGQPGWRFGNALTVLDVNGOKLTDVAIGAPGEDNRGAYVLF 540
 DB 496 PRGQVQVQCDVLYRGEGQHPGRFGAALTVLGVDNEDKLDIDVAIGAPGQENKAYVLF 555
 QY 541 HGTSGSIGSPSHSQRISPRLOVFGQSLSGGQDLTMDGLVLDLTVGAGQHVLRLSRQ 600
 DB 556 HGAESGIGSPSHSQRISPRLOVFGQSLSGGQDLTMDGLVLDLTVGAGQHVLRLSRQ 615
 QY 601 PVLRYKAIMFNPREVARNVPCNDQVQVKGKEGVRVCLHVQKSTDRLEBGOIQSVVT 660
 DB 616 PVLKVGVMRFPSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLDQL-GDIQSSVR 673
 QY 661 YDLALDSGRPHSRAPVNETKNSTRQTVGLTCTCTLKLQLPNCIEDPVSPIVLRINF 720
 DB 674 FDLALDPORLTSRAIFNETKNPTLTKRTKTLGLGHCETLKLALLPDCVEDVSPILHLNF 733
 QY 721 SLVGTPLSAFENLAPVLAEDAQRFTALFPFKKNCNDNICODDLSTITFSMSLDCLVVG 780
 DB 734 SLVREPIPSQNLRLPVLAVGSDQLFTASLPFKKNCQDGLCESGLVTLSPSGLOTITVG 793
 QY 781 GPREFNTVTVRNDGEDSYRTQVTFPPPLDLSVEKVSFTLQNRQSRQSWRLACESASSTEV 840
 DB 794 SSELNVIVTVNAGEDSYRTQVTFPPPLDLSVEKVSFTLQNRQSRQSWRLACESASSTEV 852
 QY 841 SGALKSTSCSINHPIPPENSEVNTITFDVDSKASLGNKLLKANVTSENNKPTNKTEF 900
 DB 853 EG-LRSSRCSVNHPIFHEGSGNGTPIVTFVDSYKATLADRMIMRASSENKSKSSKATF 911
 QY 901 QLELPVAVVAVVTVSHGVSTKYNLF-TASNTSRVMOHOVQVSNLQSRSLPISLVLPV 959
 DB 912 QLELPVAVVAVVTVSHGVSTKYNLF-TASNTSRVMOHOVQVSNLQSRSLPISLVLPV 971
 QY 960 VRLNQTVIWDPRQVTFSENLSSTCHTRELPSHSDPLAELKAPVAVNCISIANVQRIQDI 1019
 DB 972 VLLNGVAVVAVVTVSHGVSTKYNLF-TASNTSRVMOHOVQVSNLQSRSLPISLVLPV 1029
 QY 1020 PFGIOQBEFNATLKNLSFDWIKTSHNLLIVSTABILENDSVPTLLPGOGAFVRSQTE 1079
 DB 1030 PFSVQBEELDFTLKGNLSFGWVRETQKVLVWSVAEITDTSVYSLQPGQAFWRAQNE 1089
 QY 1080 TKVEPFEVNPPLIVGSSVGGILLALITAAALYKLGFFKQYKDMSE 1128
 DB 1090 MVLBEDEVYNAIPIIMGSSVGCALLLALITATLYKLGFFKQYKDMSE 1138
 RESULT 14
 AAQ23049
 ID AAQ23049 standard; protein; 1161 AA.
 AC AAQ23049;
 DT 24-FEB-1998 (first entry)
 XX Human beta 2 integrin alpha d subunit.
 XX Beta 2 integrin alpha d subunit; human; cell migration; cell adhesion;
 XX phagocytosis; diabetes; atherosclerosis; multiple sclerosis; asthma;
 XX psoriasis; lung inflammation; acute respiratory distress syndrome;
 XX rheumatoid arthritis.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Domain 17. .1108

```

FT      /label= Extracellular_domain
FT      150..352
FT      /note= "region homologous to the I (insertion) domain
FT      common to CD11a, CD11b and CD11c"
FT      Domain
FT      1109..1128
FT      /label= Transmembrane_domain
FT      /note= "homologous to the human CD11c transmembrane
FT      region"
FT      1129..1161
FT      /label= Cytoplasmic_domain
FT      WO9731099-A1.
XX      24-FEB-1997; 9TWO-US002713.
XX      28-AUG-1997.
XX      24-FEB-1997; 9TWO-US002713.
XX      22-FEB-1996; 9GUS-00605672.
XX      (ICOS-) ICOS CORP.
XX      Gallatin WM, van Der Vieren M;
XX      WPI; 1997-435154/40.
XX      N-PSDB; AAT79220.
XX      Hydrinoma 199M and antibody secreted by it - specific for new rat beta2
XX      integrin subunit, useful to detect subunit in cells and modulate its
XX      activity.
XX      Example 5; Page 116-120; 222pp; English.
XX      This polypeptide comprises a novel human beta 2 integrin subunit,
XX      designated alpha d. Its sequence was deduced from a cDNA clone (see
XX      AAT79220) isolated from a spleen cDNA library. Alpha d is involved in
XX      cell migration, phagocytosis and cell-cell interaction. Recombinant alpha
XX      d polypeptides can be expressed in transformed host cells for use in
XX      assays for identifying antibodies or other compounds that modulate alpha
XX      d activity or which modulate the interaction between alpha d and a
XX      ligand, for treating or preventing diseases in which macrophages are
XX      implicated. Treatment is applicable to disease states in which alpha d
XX      binding, or localised accumulation of cells which express alpha d, is
XX      implicated such as such as type I diabetes, atherosclerosis, multiple
XX      sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
XX      distress syndrome and rheumatoid arthritis
XX      Sequence 1161 AA;
XX      Query Match 57.8%; Score 3401; DB 2; Length 1161;
XX      Best Local Similarity 59.3%; Pred. No. 1e-275;
XX      Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;
XX      1 FNLDTENAMTFQENARGFGQSVVQLQGRVVGAPQEIIVAAQNGSLYQCYSTGSCPEI 60
XX      17 FNLDVEEPTIFQEDAGFGQSVVQFGGSLVVGAPLEVVAAQTCRLYDCAATGMCQPI 76
XX      61 RLQVPVAVNMSIGLSLAATTSPPLACGPTVHQTCSNTVYKGLCELPQSNLRQOQPK 120
XX      77 PLHIRPEAVNMSIGLTAAASTNGSLRLLACGPTLHVRVCGENSYSGKSCLLLSRW-BIIQT 135
XX      121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEWVSTVMBQLKSKTLFSLMOYSEEF 180
XX      136 VPDATPECHQEMDIFVLIDGSGSIDQDNDFNQMGFVQAVMQFEGTDTLFLALMYSNLL 195
XX      181 RHPTPKQNNPNRSLIKPTQLLGRTHATGKRVRELFTNITNGARKNAFKILILL 240
XX      196 KIHFTTQRTSPQSSLDVPIVQLKGLTFTATGILTVVTQLFHKGARKSAKKILIVI 255
XX      241 TDEKFGDPLGVDYVPELDRCGVTRYVIGVDARFSEKSRQELNVTASKPRDRHVFQIN 300
XX      256 TDQCKYKDFLEYSVDVLPQEKAGIIRYALGVGHACGPTARQELNTISSAPPQDHVKVD 315
XX      301 NFPEALXTIQNLEKIFALTEGTQGTGSSSFHEHMSQEGFSAATISNGPLLSTGVSYDNAG 360

```

```

Db      316 NFPAALASIQKQOEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMGDLFGAVGSFWSG 375
Qy      361 GVFLYTSKEKSTFINMTRVDSMDNDAYLYAAAIILRNVRQSVLVGAPRVOHIGLVAMFR 420
Db      376 GAFLYPNMSPPTINMSQENVDNRDSYLGSTELALWKGQVNLVGLAPRYQHTGKAVIFT 435
Qy      421 QNTGMVESANVKGTQIGAFYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGQGVSCPL 480
Db      436 QVSRQWRKKAETVGTQIGSYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGQGVSCPL 495
Qy      481 PRQARAWQCDAYLYGEGQGPWGRFGAALTVLGVNNGDKLTQVAIGAPGEDNRGAYVLF 540
Db      496 PRGORVQWQCDAYLRGEGQHPWGRFGAALTVLGVNNGDKLTQVAIGAPGEDNRGAYVLF 555
Qy      541 HGTSGSGISPSHSORITAGSKLSPRLQYFGOSLSSGGDLTMGDLVLTGVAQGHVLLRSQ 600
Db      556 HGASESGISPSHSORITAGSKLSPRLQYFGOSLSSGGDLTMGDLVLTGVAQGHVLLRSQ 615
Qy      601 PVLRLVAIMEFNPREVARNVFECDNVVKGKEAGEVRVCLHVQKSTRDLREBQIQSVWT 660
Db      616 PVLKVGVMREFSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673
Qy      661 YDLALDSGRPHSRVAVNETKNSRROTQVULGLTQTCETLKLQLPNCIEDPVSIVLBNF 720
Db      674 FDLALDSGRPHSRVAVNETKNSRROTQVULGLTQTCETLKLQLPNCIEDPVSIVLBNF 733
Qy      721 SLVGTPLSAFGRNRPVLAEDAQRLFTALPFFKNCNDNICQDDLSLTFSEFMSLDCLVWG 780
Db      734 SLVREPIPSQNLRPVLAQVGSQDLFTASLPFFKNCNDNICQDDLSLTFSEFMSLDCLVWG 793
Qy      781 GPREFNVTVVRNDGEDSYRTQVFFPDLTLVSKYSTLQNRQSRQSRWELACESASTEV 840
Db      794 SSLELVIVTVNRAGDSYGTVSVSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 852
Qy      841 SGALKSTSCSINPIRPNSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
Db      853 EG-LRSRCVSNPIRPIHGGNGFIVTFDVSYKATLDGRLMRASSENKASSKATP 911
Qy      901 QLELPVYAVYVMVTHGVSTKYLPN-TASENTSRVWQYQVSNLQQRSLPISLVLTPV 959
Db      912 QLELPVYAVYVMVTHGVSTKYLPN-TASENTSRVWQYQVSNLQQRSLPISLVLTPV 971
Qy      960 VRLNQTVMDRPQVTPSENLSSTCHTKERLPKSHSDFLAELRKAPVNVCSIAVCORICDI 1019
Db      972 VLLNGVAVVDVVMVMEAPQSL--PCVSRKPPQSHSDFLTQISRSPMLDCSIADCLQFPCDV 1029
Qy      1020 PFGIOEFNATLKGNLSPDWYIYKTSNHLIIVSTABILFNDSVFTLLPQOGAPVRSQTE 1079
Db      1030 PSFVSQSELDFTLKGNLSPDWYIYKTSNHLIIVSTABILFNDSVFTLLPQOGAPVRSQTE 1089
Qy      1080 TKVEPPFVPHPLIVGVSSVGGILLIITAAALYKLGFEKRYKQVMSE 1128
Db      1090 MVLDEDSVYNAIPINGSSVGAALLLALITATLYKLGFFKRYKQVMSE 1138
XX      RESULT 15
XX      AAW57491
XX      ID AAW57491 standard; protein; 1161 AA.
XX      AC AAW57491;
XX      DT 24-AUG-1998 (first entry)
XX      DE Human beta2 integrin alpha subunit (alpha d) polypeptide.
XX      KW Beta 2 integrin alpha subunit; alpha d; human; treatment; diabetes;
XX      reporter-transactivator construct; arteriosclerosis; atherosclerosis;
XX      inflammatory bowel disease; arthritis; multiple sclerosis.
XX      OS Homo sapiens.
XX      XX Key Location/Qualifiers

```

```

Peptide      1..16
/Note= "signal peptide"
Protein      17..1161
/Note= "mature protein"
Domain      17..1108
/Note= "extracellular domain"
Domain      150..352
/Note= "I (insertion) domain"
Region      1109..1128
/Note= "transmembrane region"
Domain      1129..1161
/Note= "cytoplasmic domain"

US5766850-A.
16-JUN-1998.
21-DEC-1994; 94US-00362652.
23-DEC-1993; 93US-00173497.
05-AUG-1994; 94US-00286889.

XX (ICOS-) ICOS CORP.
XX Van Der Vieren M, Gallatin WM;
XX WPI; 1998-361678/31.
XX N-PSDB; RAV31540.
XX
XX Isolation of DNA encoding protein that binds to integrin subunit - using
XX recombinant cells containing reporter-transactivator construct.
XX
XX Example 5; Col 45-54; 86pp; English.
XX
XX This represents a human beta2 integrin alpha subunit (alpha d)
XX polypeptide. This is used in the methods of the invention for isolating a
XX polynucleotide encoding a protein that binds to alpha d. The method
XX comprises transforming or transfecting host cells with a DNA construct
XX comprising a reporter gene under the control of a promoter regulated by a
XX transcription factor having a DNA-binding domain and an activating
XX domain. A first hybrid DNA sequence encoding a fusion of at least part of the
XX alpha d and either the DNA-binding domain or the activating domain of the
XX transcription factor and a library of hybrid DNA sequences encoding
XX fusions of at least part of putative alpha d-binding proteins and the DNA
XX binding domain or the activating domain of the transcription factor
XX which is not incorporated in the first fusion are expressed in the host
XX cells. The binding of an alpha d-binding protein to alpha d in a
XX particular host cell is detected by determining production of the
XX reporter gene product in the cell. The hybrid DNA sequence encoding the
XX alpha d-binding protein can be isolated from the cell. Alpha d may be
XX useful for treating graft arteriosclerosis, atherosclerosis, diabetes,
XX inflammatory bowel disease, arthritis and multiple sclerosis
XX
XX Sequence 1161 AA;
XX
XX Query Match      57.8%; Score 3401; DB 2; Length 1161;
XX Best Local Similarity 59.3%; Ered. No. 1e-275;
XX Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;
XX
XX 1 FNLDTENAMTFQBNARGFGQSVVQLOGSRVVGAPQEIIVANQGRSLYQCDSYSGCEPI 60
XX 17 FNLDVBEPTTFQBDAGGFGQSVVQFGSRLVVGAPLEVVAAQTGLYDCAATGMQCP 76
XX
XX 61 RLOVPEAVNMSLGLSLAATTSPPQLACGPTVTHOTCSNTVYKGLCFELFGSNLRQPOK 120
XX 77 PLHIRPANNVSLGLHASTNSRLACGPTLHRCVGENSYSGKSLGLSRW-EIIQT 135
XX
XX 121 FPEALRGCPQSDIAFLIDGSGSIIPHDFRMKEWVSTVMEQLKSKTFLSLMOYSEEP 180
XX 136 VPDATPCRPQEMDIVFLIDGSGSIDQDNFNQKGFVQAVMGQFEGTDTLFLALMOYSNLL 195
XX
XX 181 RIHPTPEFQNNPRSLIKPITQLGRTHATGRKVVRELENTNGARKNAKILFLL 240
XX

```

```

Db 196 KIHFTFTQRTSPSQSLVDPIVQLKGLTFTATGILTVVTLFHHKNGARKSAKILLIVI 255
Qy 241 TDGEKFGDPLGYEDVPIBELDREGVIRYVIVGSDAPRSEKSRQELMTVASKPRDHVFOIN 300
Db 256 TDGQKYKDPLEYSDVIPAERKAGIIRYAIYGVGHAFQGPATQELMTISSAPPQDFVFD 315
Qy 301 NPEALKTIQNLREKIFAIEGTQTGSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 316 NFAALGSIQKLOEKIYAVEGTQSRASSSPQHEMSQEGFSTALTMDGLFLGAVGSFWSG 375
Qy 361 GVPFLYTSKESKSTFINWTRVDSMDNDAYLGYAAAIIILNRVQSLVILGAPRYOHIGLVAMFR 420
Db 376 GAFLYFPNMSPTFINMSQENVDMSYLGYSYELATWGVQVQLVILGAPRYOHTKAVIET 435
Qy 421 QNTGMWESNANYKGTQIGAYFCASLCSVDVDSNGSTDLVLIGABHYTEQTRGGQSVCP 480
Db 436 QVSRQWRKKAETVGTQIGSYFCASLCSVDVDSNGSTDLVLIGABHYTEQTRGGQSVCP 495
Qy 481 PRGORARWOCDAVLYGEOGPMGRGALTVLGVNKGKLTQVAGTAPGREDNEGAYLYF 540
Db 496 PRGORVQWOCDAVLYGEOGPMGRGALTVLGVNKGKLTQVAGTAPGREDNEGAYLYF 555
Qy 541 HGTSGSIGSPSHSQRIAGSKLSPRLQYFQSLSGQDLTMDGLVDLTGAGQHVLRLRSQ 600
Db 556 HGASESGISPSHSQRIASSQLSPRLQYFQSLSGQDLTMDGLVDLTGAGQHVLRLRSQ 615
Qy 601 PVLRYKAIMENPREVARNVFECDQVVKGEAGEVRYVCLHVOKSTRDRLEGOIQSVVT 660
Db 616 PVLKGVAMRFPVEVAKAVYRCHEEKPSALRAGDATVCLTIQKSSLDOL--GDIQSVR 673
Qy 661 YDLALDSGPHGRAVFNETKSTRQTVLGTOTCTFLKLQLENCIEDPVSPIVLRINF 720
Db 674 FDLALDPGLTSLRAIFNETKTPTRRTKLTGLGHICETKLLKLLPCDVEDVVSPIILHNF 733
Qy 721 SLVGPPLSAFGLRPLVAEDAQRFLTALFPPEKKNCGNDNICODDLISITFSFMSLDCLVVG 780
Db 734 SLVREPLSPQNLREPLVAVGSDLTASLPFEKKNCGDGLCEGDLGVTLSFSGLTGTVG 793
Qy 781 GPREFNVTYVNDGEDSYRQVTFPPPLDLSYRKVSTLQNRQSRWRWLACBSASTEV 840
Db 794 SLELNVIYVWVNDGEDSYGTWVSLYFAGLSHRVSGAQPHQSAALRLACETV-PTED 852
Qy 841 SGALKSTCSINHPIFPENSEVTFNITPDVDSKASLGNKLLKANVTSENMMPTNKTET 900
Db 853 EG-LRSSRCSVNHPIFHBGSGNGTFTVITFDVSYKATLGDRLMRASASSENKASSKATP 911
Qy 901 QLELPKYAVVWVTSHGVSYTKYLP-TASENTRVWQHOYQVSNLQORSILPISLVLYP 959
Db 912 QLELPKYAVVWVTSHGVSYTKYLP-TASENTRVWQHOYQVSNLQORSILPISLVLYP 971
Qy 960 VRINQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVAVNCIAVCORICDI 1019
Db 972 VLLNGVAVWVWMEAPSQSL--PCVSEKRPPOHSDFLTQISRSPLMDCSIADCLQPRCDV 1029
Qy 1020 PFGTOEENATLKNLSFDWYIKTSHNHLIVSTAELFENDSVFTLLPGOGAFVRSOTE 1079
Db 1030 PSFSVQOEELDFTLKRLNLSFGVRETLOCKVLVWSVAEITFDTSVYSQLPQOEAFWRAQME 1089
Qy 1080 TKVEPFEVNPFLPLVSGSVGGLLALLITAALYKLGFFKRYQKDMSE 1128
Db 1090 MYLEDEEVNVAIPIIMGSSVGALLLALLITATLYKLGFFKRYKHEMLED 1138

```

Search completed: June 7, 2004, 17:11:39
Job time : 53.4608 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:16:51 ; Search time 38.8905 Seconds

(without alignments)
8225.189 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTFOENARGFGQ.....FKROYKDMWSEGGPGABPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5884	100.0	1137	10	US-09-902-481A-4
2	5857	99.5	1137	10	US-09-902-481A-3
3	5852	99.5	1153	9	US-09-350-259-3
4	5852	99.5	1153	10	US-09-902-481A-1
5	5852	99.5	1153	10	US-09-891-943-3
6	5852	99.5	1153	14	US-10-144-259-30
7	5852	99.5	1153	14	US-10-207-655-176
8	5851	99.4	1137	10	US-09-902-481A-5
9	5845	99.3	1137	10	US-09-902-481A-6
10	5836.5	99.2	1152	9	US-09-945-265-4
11	3459	58.8	1163	15	US-10-116-275-204
12	3436	58.4	1163	9	US-09-350-259-4
13	3436	58.4	1163	10	US-09-891-943-4
14	3401	57.8	1161	9	US-09-350-259-2
15	3401	57.8	1161	10	US-09-891-943-2

16 3385.5 57.5 1161 9 US-09-350-259-99
17 3385.5 57.5 1161 10 US-09-891-943-99
18 3224.5 54.8 1161 9 US-09-350-259-55
19 3224.5 54.8 1161 10 US-09-891-943-55
20 3212 54.6 1161 9 US-09-350-259-53
21 3212 54.6 1161 10 US-09-891-943-53
22 3203.5 54.4 1151 9 US-09-350-259-37
23 3203.5 54.4 1151 10 US-09-891-943-37
24 3191 54.2 1155 9 US-09-350-259-46
25 3191 54.2 1155 10 US-09-891-943-46
26 1848 31.4 369 12 US-10-087-192-1212
27 1547.5 26.3 1170 9 US-09-945-265-2
28 1543.5 26.2 1170 12 US-10-361-164-1
29 1520 25.8 1223 16 US-10-408-765A-295
30 1350.5 23.0 1086 16 US-10-408-765A-1871
31 1229.5 20.9 494 9 US-09-350-259-103
32 1229.5 20.9 494 10 US-09-891-943-103
33 1141.5 19.4 413 9 US-09-350-259-101
34 1141.5 19.4 413 10 US-09-891-943-101
35 1140 19.4 1179 14 US-10-177-550-2
36 1140 19.4 1179 15 US-10-173-551-2
37 1093.5 18.6 1151 10 US-09-894-130-103
38 1093.5 18.6 1151 10 US-09-836-353A-103
39 1093.5 18.6 1179 12 US-09-318-715-250
40 1092 18.6 1188 15 US-10-291-265-338
41 1088 18.5 1188 15 US-10-291-265-338
42 1084.5 18.4 1189 10 US-09-894-130-35
43 1084.5 18.4 1189 10 US-09-836-353A-35
44 1084.5 18.4 1189 12 US-10-362-839-4
45 1074 18.3 589 12 US-10-361-164-2

ALIGNMENTS

RESULT 1

US-09-902-481A-4

; Sequence 4, Application US/09902481A

; Publication No. US2003005440A1

; GENERAL INFORMATION:

; APPLICANT: Springer, Timothy

; APPLICANT: Shimaoka, Motomu

; APPLICANT: Shifman, Julia

; APPLICANT: Mayo, Stephen

; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY

; FILE REFERENCE: A-70586-1/RFT/RMS/RMK

; CURRENT APPLICATION NUMBER: US/09/902.481A

; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/216,600

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 1137

; TYPE: PRT

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: synthetic

US-09-902-481A-4

Query Match 100.0%; Score 5884; DB 10; Length 1137;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGQSVVVGAPQIIVANQSGSLVQCDYSTGSCPEI 60

Db 1 FNLDTENAMTFOENARGFGQSVVVGAPQIIVANQSGSLVQCDYSTGSCPEI 60

Qy 61 RLQVPVAVNMVSLGLSLAAATTPPQLLACGPTVHTQTCSENTYVKGCLFPLFQSNLRQQPQK 120

Db 61 RLQVPVAVNMVSLGLSLAAATTPPQLLACGPTVHTQTCSENTYVKGCLFPLFQSNLRQQPQK 120

Qy 121 FPEALRGCPQEDSDIAFLIDGSGIIPHDFRMKEWSTVWEQLKSKTLPFLWQYSSEF 180

121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKEWSTVMEQLKSKTLFSLMQYSSEF 180
181 RHFTFKFQNNPNPSLSIKPITQLLGRTHATGLKRVVRELFTNITNGARKNAFKILFLL 240
181 RHFTFKFQNNPNPSLSIKPITQLLGRTHATGLKRVVRELFTNITNGARKNAFKILFLL 240
241 TDEKFGDPLGYEDVPELDREGVIRVYLVGDAPFSEKSRQBLNTVASKPRDHVFPQIN 300
241 TDEKFGDPLGYEDVPELDREGVIRVYLVGDAPFSEKSRQBLNTVASKPRDHVFPQIN 300
301 NFEALKTIONQUREKIPAIETGTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
301 NFEALKTIONQUREKIPAIETGTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
361 GVFLYTSKEKSTFIMNTRVDSMDNDAYLGVAAILLRNVQSLVLCAPRYQHIGLVAMFR 420
361 GVFLYTSKEKSTFIMNTRVDSMDNDAYLGVAAILLRNVQSLVLCAPRYQHIGLVAMFR 420
421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSVCP 480
421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSVCP 480
481 PRGQARWQCDVLYGEOGQOPGRFGAALTVDVNGDXLTDVAIGAPGEEDNRGAVYLP 540
481 PRGQARWQCDVLYGEOGQOPGRFGAALTVDVNGDXLTDVAIGAPGEEDNRGAVYLP 540
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
601 PVLKVKAIEMFNPREVARNVPECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
601 PVLKVKAIEMFNPREVARNVPECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
661 YDLALDSGRPHSRVAFVNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
661 YDLALDSGRPHSRVAFVNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
721 SLVGTPLSAFGLNRPVLAEDAQRFTALPFPKKNCGNDNICQDDLSITTFSPMSLCLVVG 780
721 SLVGTPLSAFGLNRPVLAEDAQRFTALPFPKKNCGNDNICQDDLSITTFSPMSLCLVVG 780
781 GRPEFNVTVTVNDGSDVYRTQVTPFPPLDLSVRKYSTLQNRORSKRWLACESASSTEV 840
781 GRPEFNVTVTVNDGSDVYRTQVTPFPPLDLSVRKYSTLQNRORSKRWLACESASSTEV 840
841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
901 QLELPVKYAVMVVTSHGVSSTKYLNTFASNTSRVMOHQYQVSNLQORSPLSLVFLVPV 960
901 QLELPVKYAVMVVTSHGVSSTKYLNTFASNTSRVMOHQYQVSNLQORSPLSLVFLVPV 960
961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDPLAELRKAPVWNCISIAVQCIQCDIP 1020
961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDPLAELRKAPVWNCISIAVQCIQCDIP 1020
1021 PFGIOBEFNATLKNLSFOWYIKTSHNLLIYVSTABIFLNDVSPVLLPQOGAFVRSOTET 1080
1021 PFGIOBEFNATLKNLSFOWYIKTSHNLLIYVSTABIFLNDVSPVLLPQOGAFVRSOTET 1080
1081 KYEPPFVNPPLPLIVGSSVGGILLALITAAALYKLGFFKRYQKDMWSEGGPPGABPQ 1137
1081 KYEPPFVNPPLPLIVGSSVGGILLALITAAALYKLGFFKRYQKDMWSEGGPPGABPQ 1137

RESULT 2

US-09-902-481A-3
; Sequence 3, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:

; APPLICANT: Springer, Timothy
; APPLICANT: Shimooka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
; US-09-902-481A-3

Query Match 99.5%; Score 5857; DB 10; Length 1137;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1131; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 FNLDTENAMTQENARFGQSVQIQSGRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
DB 1 FNLDTENAMTQENARFGQSVQIQSGRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
QY 61 RLQVPVEAVNNSLGLSLAATTPPQILLACGPTVHTQTCENTYVKGCLPFGSNLRQQPQK 120
DB 61 RLQVPVEAVNNSLGLSLAATTPPQILLACGPTVHTQTCENTYVKGCLPFGSNLRQQPQK 120
QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKEWSTVMEQLKSKTLFSLMQYSSEF 180
DB 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKEWSTVMEQLKSKTLFSLMQYSSEF 180
QY 181 RHFTFKFQNNPNPSLSIKPITQLLGRTHATGLKRVVRELFTNITNGARKNAFKILFLL 240
DB 181 RHFTFKFQNNPNPSLSIKPITQLLGRTHATGLKRVVRELFTNITNGARKNAFKILFLL 240
QY 241 TDGSKGDPGLGYEDVPELDREGVIRVYLVGDAPFSEKSRQBLNTVASKPRDHVFPQIN 300
DB 241 TDGSKGDPGLGYEDVPELDREGVIRVYLVGDAPFSEKSRQBLNTVASKPRDHVFPQIN 300
QY 301 NFEALKTIONQUREKIPAIETGTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 301 NFEALKTIONQUREKIPAIETGTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
QY 361 GVFLYTSKEKSTFIMNTRVDSMDNDAYLGVAAILLRNVQSLVLCAPRYQHIGLVAMFR 420
DB 361 GVFLYTSKEKSTFIMNTRVDSMDNDAYLGVAAILLRNVQSLVLCAPRYQHIGLVAMFR 420
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSVCP 480
DB 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSVCP 480
QY 481 PRGQARWQCDVLYGEOGQOPGRFGAALTVDVNGDXLTDVAIGAPGEEDNRGAVYLP 540
DB 481 PRGQARWQCDVLYGEOGQOPGRFGAALTVDVNGDXLTDVAIGAPGEEDNRGAVYLP 540
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
DB 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
QY 601 PVLKVKAIEMFNPREVARNVPECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 601 PVLKVKAIEMFNPREVARNVPECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
QY 661 YDLALDSGRPHSRVAFVNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
DB 661 YDLALDSGRPHSRVAFVNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
QY 721 SLVGTPLSAFGLNRPVLAEDAQRFTALPFPKKNCGNDNICQDDLSITTFSPMSLCLVVG 780

Db 721 SLVGTPLSAFGLNLRVLAEDAQRLLTALFPPEKNCNDNI QODDLSITFSFMSLDCLVVG 780
Qy 781 GPREFNVTVTVRNDEGDSYRQVTFPPFLDLSYRKVSTLQNSORSWRWLACESASSTEV 840
Db 781 GPREFNVTVTVRNDEGDSYRQVTFPPFLDLSYRKVSTLQNSORSWRWLACESASSTEV 840
Qy 841 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 900
Db 841 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 900
Qy 901 QLELPVKAVYVMVTSYHGVSTKYLNTASNTSRVMQHOYQVSNLQORSLSISLVLVVPV 960
Db 901 QLELPVKAVYVMVTSYHGVSTKYLNTASNTSRVMQHOYQVSNLQORSLSISLVLVVPV 960
Qy 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP 1020
Db 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP 1020
Qy 1021 FFGIOEEFNATLKGNSLSDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1080
Db 1021 FFGIOEEFNATLKGNSLSDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1080
Qy 1081 KVEPPEVFNPLPLIVGSSVGGILLALITAAALYKLGFFKRYQKDMWSEGGPPGAEPO 1137
Db 1081 KVEPPEVFNPLPLIVGSSVGGILLALITAAALYKLGFFKRYQKDMWSEGGPPGAEPO 1137

RESULT 3
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.5%; Score 5852; DB 9; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
Qy 1 FNLDTENATFOENARGFGQSVVQLQSGRVVVGAPQEI VAAANQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENATFOENARGFGQSVVQLQSGRVVVGAPQEI VAAANQSGSLYQCDYSTGSCPEI 76
Qy 61 RLOVPVEAVNNSLGLSLAATSPOLLACGPTVHTQCSNTYVVKGLCFPLGSLNRQQPOK 120
Db 77 RLOVPVEAVNNSLGLSLAATSPOLLACGPTVHTQCSNTYVVKGLCFPLGSLNRQQPOK 136
Qy 121 FPEALRGCPQEDSDIAFLIDSGSIIIPHDPRMKKSWTVMQOLKSKTLPFLSMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDSGSIIIPHDPRMKKSWTVMQOLKSKTLPFLSMQYSEEF 196

Qy 181 RIHFTPFKEFQNNPNRSLIKPITQLLGRTHATGLRKVVRLEFNITNGARKNAFKILFLL 240
Db 197 RIHFTPFKEFQNNPNRSLIKPITQLLGRTHATGLRKVVRLEFNITNGARKNAFKILFLL 256
Qy 241 TDEKPKDPLGYEDVPELDRREGVIRVYGVGDAFRSEKSRQELNLTVASPPRDRHVQIN 300
Db 257 TDEKPKDPLGYEDVPELDRREGVIRVYGVGDAFRSEKSRQELNLTVASPPRDRHVQIN 316
Qy 301 NFEALKTIQNLREKIPAIETGTGSSSSPEHEMSQEGFSAATNSNGPLLSITVGSYDWA 360
Db 317 NFEALKTIQNLREKIPAIETGTGSSSSPEHEMSQEGFSAATNSNGPLLSITVGSYDWA 376
Qy 361 GVFLYTSKESKSTFNIMTRVDSNDNDVILGYAAAIILRNRVQSLVGLAPRVOHIGLVAMER 420
Db 377 GVFLYTSKESKSTFNIMTRVDSNDNDVILGYAAAIILRNRVQSLVGLAPRVOHIGLVAMER 436
Qy 421 QNTGMESSNANVKGTOIGAYFGASLSCSDVDNSGSDTLVLIGAPHYTYEQTRGGQVSCPL 480
Db 437 QNTGMESSNANVKGTOIGAYFGASLSCSDVDNSGSDTLVLIGAPHYTYEQTRGGQVSCPL 496
Qy 481 PRGRARWQCDVILYGEQGFWRFGAALTVLGDVNGDKLTDVAIGAPGEDNKGAYLIF 540
Db 497 PRGRARWQCDVILYGEQGFWRFGAALTVLGDVNGDKLTDVAIGAPGEDNKGAYLIF 556
Qy 541 HGTSGSGISPSHQRISAGSKLSPRLQYFGQSLSGQODLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHQRISAGSKLSPRLQYFGQSLSGQODLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMBENPREVARNVFECNDQVQVKEAGEVAVCLHVQKSTDRDLREGQIQSVVT 660
Db 617 PVLRVKAIMBENPREVARNVFECNDQVQVKEAGEVAVCLHVQKSTDRDLREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAVFNETKNSTRTOVLGLTQCTETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSRVAVFNETKNSTRTOVLGLTQCTETLKLQLPNCIEDPVSPIVLRNLF 736
Qy 721 SLVGTPLSAFGLNLRVLAEDAQRLLTALFPPEKNCNDNI QODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLRVLAEDAQRLLTALFPPEKNCNDNI QODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVTVRNDEGDSYRQVTFPPFLDLSYRKVSTLQNSORSWRWLACESASSTEV 840
Db 797 GPREFNVTVTVRNDEGDSYRQVTFPPFLDLSYRKVSTLQNSORSWRWLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 900
Db 857 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 916
Qy 901 QLELPVKAVYVMVTSYHGVSTKYLNTASNTSRVMQHOYQVSNLQORSLSISLVLVVPV 960
Db 917 QLELPVKAVYVMVTSYHGVSTKYLNTASNTSRVMQHOYQVSNLQORSLSISLVLVVPV 976
Qy 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP 1036
Qy 1021 FFGIOEEFNATLKGNSLSDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1080
Db 1037 FFGIOEEFNATLKGNSLSDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1096
Qy 1081 KVEPPEVFNPLPLIVGSSVGGILLALITAAALYKLGFFKRYQKDMWSEGGPPGAEPO 1137
Db 1097 KVEPPEVFNPLPLIVGSSVGGILLALITAAALYKLGFFKRYQKDMWSEGGPPGAEPO 1153

RESULT 4
US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia

APPLICANT: Mayo, Stephen
TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
FILE REFERENCE: A-70586-1/FT/MS/RMK
CURRENT APPLICATION NUMBER: US/09/902,481A
CURRENT FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/216,600
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patent in version 3.1
SEQ ID NO 1
LENGTH: 1153
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: mat peptide
LOCATION: (17)..{}
OTHER INFORMATION:
US-09-902-481A-1

Query Match 99.5%; Score 5852; DB 10; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFQSVVQLQGSRRVVGAPQEIIVAAANQSGSLVQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFQSVVQLQGSRRVVGAPQEIIVAAANQSGSLVQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHCTSENTYVKGCLFGLFSGNLRQQPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHCTSENTYVKGCLFGLFSGNLRQQPQK 136
QY 121 FPEALRGCPQSDIAFLIDGSGIIPHDPRFMKEVSTVMEQKKSLTFLSLMOYSEF 180
DB 137 FPEALRGCPQSDIAFLIDGSGIIPHDPRFMKEVSTVMEQKKSLTFLSLMOYSEF 196
QY 181 RHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 240
DB 197 RHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 256
QY 241 TGEKFGDPLGVEDVITPELDREGVIRYVIGVDGAPRSEKROELNTVASKPPRDRHVQIN 300
DB 257 TGEKFGDPLGVEDVITPELDREGVIRYVIGVDGAPRSEKROELNTVASKPPRDRHVQIN 316
QY 301 NFEALATIONLREKIFALEGTQTGSSSFHEHMSQEGFSAITNSGPLLSTVGVSDWAG 360
DB 317 NFEALATIONLREKIFALEGTQTGSSSFHEHMSQEGFSAITNSGPLLSTVGVSDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAYAAAILLRNVQSLVLGAPRYQHIGLVAMPR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAYAAAILLRNVQSLVLGAPRYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLLVLGAPHYVECTRGQVSVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLLVLGAPHYVECTRGQVSVCP 496
QY 481 PRQORARWQCDVLYGEGQOPWCFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRQORARWQCDVLYGEGQOPWCFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPHSQRIAGSKLSPRLQYFGOSLGGQDLTWDGLVDLTVCAGQHVLLRSQ 600
DB 557 HGTSGSGISPHSQRIAGSKLSPRLQYFGOSLGGQDLTWDGLVDLTVCAGQHVLLRSQ 616
QY 601 PVLRVKAIMEFNPREVARNVFCNDQVVKGEAGVRVCLHVOKSTRDLRREGQIQSVVT 660
DB 617 PVLRVKAIMEFNPREVARNVFCNDQVVKGEAGVRVCLHVOKSTRDLRREGQIQSVVT 676
QY 661 YDLALDSGRPHSAVFNETKQSTRQTVLGLTQTCETLKLQLPNCIEDPVPSPVLRINF 720
DB 677 YDLALDSGRPHSAVFNETKQSTRQTVLGLTQTCETLKLQLPNCIEDPVPSPVLRINF 736
QY 721 SLVGTPLSARGNLRPVLAEDAQRFLTALFPFKNCGNDNICQDDLSITFSFMSLDCLVVG 780

DB 737 SLVGTPLSARGNLRPVLAEDAQRFLTALFPFKNCGNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGDSYRTQVTFPEPLDLSYRKVSTLQNSORSRWRLACSSASSTEV 840
DB 797 GPREFNVTVVRNDGDSYRTQVTFPEPLDLSYRKVSTLQNSORSRWRLACSSASSTEV 856
QY 841 SGALKSTCSINHPFIPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 900
DB 857 SGALKSTCSINHPFIPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 916
QY 901 QLELPVKYAVYVMVTSKGVSTKYNFTASENTSRVMOHOYVSNLQORSILPISLVLFPV 960
DB 917 QLELPVKYAVYVMVTSKGVSTKYNFTASENTSRVMOHOYVSNLQORSILPISLVLFPV 976
QY 961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWVNCIAVCQRIQCDIP 1020
DB 977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWVNCIAVCQRIQCDIP 1036
QY 1021 PFGIOBEFNATLKNLSFDWYIKTSHNHLIVSTABILFNDVSFTLLPGOGAFVRSOTET 1080
DB 1037 PFGIOBEFNATLKNLSFDWYIKTSHNHLIVSTABILFNDVSFTLLPGOGAFVRSOTET 1096
QY 1081 KVPEPEVPNPPLPLVGVSSVGGLLILALITAAALYKLGFFKQYKDMNSEGGPPGAEPO 1137
DB 1097 KVPEPEVPNPPLPLVGVSSVGGLLILALITAAALYKLGFFKQYKDMNSEGGPPGAEPO 1153
RESULT 5
US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. US20030077278A1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-3

Query Match 99.5%; Score 5852; DB 10; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFQENARGFQSVVQLQGSRRVVGAPQEIIVAAANQSGSLVQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFQSVVQLQGSRRVVGAPQEIIVAAANQSGSLVQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHCTSENTYVKGCLFGLFSGNLRQQPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHCTSENTYVKGCLFGLFSGNLRQQPQK 136
QY 121 FPEALRGCPQSDIAFLIDGSGIIPHDPRFMKEVSTVMEQKKSLTFLSLMOYSEF 180
DB 137 FPEALRGCPQSDIAFLIDGSGIIPHDPRFMKEVSTVMEQKKSLTFLSLMOYSEF 196
QY 181 RHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 240
DB 197 RHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 256

```
241 TDGEKFGDPLGYEDVIVPELDREGVIRVIVGVDGAFRSEKSRQELNVTASXPPDRHVPQIN 300
Db TDGEKFGDPLGYEDVIVPELDREGVIRVIVGVDGAFRSEKSRQELNVTASXPPDRHVPQIN 316
301 NFEALKTIONLRKIKFAIEGTOTGSSSPHEMSQGFSAATISNGPLLSITVGSYDWAG 360
Db NFEALKTIONLRKIKFAIEGTOTGSSSPHEMSQGFSAATISNGPLLSITVGSYDWAG 376
361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436
421 QNTGWSNANVKGTQIAYFGASLCSVDVDSNGSTDLVIGAPHYEOTRGGOVSCPL 480
Db QNTGWSNANVKGTQIAYFGASLCSVDVDSNGSTDLVIGAPHYEOTRGGOVSCPL 496
481 PRGORARWQCDVAVLYGEOQFPGWGFAGALTVDVGVNGDKLTDVAGAPGEEDNRGAVYLF 540
Db PRGORARWQCDVAVLYGEOQFPGWGFAGALTVDVGVNGDKLTDVAGAPGEEDNRGAVYLF 556
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
601 PVLRVKAIMBNPREVARNVFECDQVVKKEAGEVRVCLHVQKSTRDRLEGGQIQSVVT 660
Db PVLRVKAIMBNPREVARNVFECDQVVKKEAGEVRVCLHVQKSTRDRLEGGQIQSVVT 676
676 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
736 SLVGTPLSAFCNLRPVLAEDAQRLLFTALFPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 780
Db SLVGTPLSAFCNLRPVLAEDAQRLLFTALFPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 796
796 GPREFNVTYVRNDGEDSYRTQVTFPPFDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db GPREFNVTYVRNDGEDSYRTQVTFPPFDLSYRKVSTLQNRQSRWRLACESASSTEV 856
856 SGALSTSCSINHPFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
Db SGALSTSCSINHPFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916
916 QLELPVKAVYVMTSHGVSTKYLNTASNTSRVWQHQQYQVSNLQSRSLPISLVFLVPV 960
Db QLELPVKAVYVMTSHGVSTKYLNTASNTSRVWQHQQYQVSNLQSRSLPISLVFLVPV 976
976 RLNQTVINDRPOVTFSENLSSTCHTKERLPSHSDFLAHLRKAIPVNCIAVCORIQCIP 1020
Db RLNQTVINDRPOVTFSENLSSTCHTKERLPSHSDFLAHLRKAIPVNCIAVCORIQCIP 1036
1036 PFGIOEFNATLKGNLSPDWYIKTSHNLLIVSTABILLFNDVFTLLPQCGAFVRSQTET 1080
Db PFGIOEFNATLKGNLSPDWYIKTSHNLLIVSTABILLFNDVFTLLPQCGAFVRSQTET 1096
1096 KVEPPEVPNPLPLIVGSSVGGILLALITAAALYKLGFFKRYQKDMWSEGGPPGABPQ 1137
Db KVEPPEVPNPLPLIVGSSVGGILLALITAAALYKLGFFKRYQKDMWSEGGPPGABPQ 1153
```

RESULT 6

```
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
```

```
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-144-259-30
```

```
Query Match 99.5%; Score 5852; DB 14; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 FNLDENAMTFQENARFGQSVVQLQSSRVVWGAPOEIVAAQNGSGSYLQCDYSTGSCPEI 60
Db 17 FNLDENAMTFQENARFGQSVVQLQSSRVVWGAPOEIVAAQNGSGSYLQCDYSTGSCPEI 76
61 RLOVPVEAVNMSLGLSLAATTSPPQLLACQPTVHQTCSNTYVVKGLCFLLFGSNLRQOPQK 120
Db 77 RLOVPVEAVNMSLGLSLAATTSPPQLLACQPTVHQTCSNTYVVKGLCFLLFGSNLRQOPQK 136
121 FPBALRGCPQEDSDIAFLIDGSGSIIPHPDRPKWKEVSTWEOQKKSKTLFSLMOYSEEF 180
Db 137 FPBALRGCPQEDSDIAFLIDGSGSIIPHPDRPKWKEVSTWEOQKKSKTLFSLMOYSEEF 196
181 RIHPTKEFQONNPNPSLIKPIITQLGRHTATGLRAVVELEFNI TNGARKNAFKILFLL 240
Db 197 RIHPTKEFQONNPNPSLIKPIITQLGRHTATGLRAVVELEFNI TNGARKNAFKILVVI 256
241 TDGEKFGDPLGYEDVIVPELDREGVIRVIVGVDGAFRSEKSRQELNVTASXPPDRHVPQIN 300
Db 257 TDGEKFGDPLGYEDVIVPELDREGVIRVIVGVDGAFRSEKSRQELNVTASXPPDRHVPQIN 316
301 NFEALKTIONLRKIKFAIEGTOTGSSSPHEMSQGFSAATISNGPLLSITVGSYDWAG 360
Db 317 NFEALKTIONLRKIKFAIEGTOTGSSSPHEMSQGFSAATISNGPLLSITVGSYDWAG 376
361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436
421 QNTGWSNANVKGTQIAYFGASLCSVDVDSNGSTDLVIGAPHYEOTRGGOVSCPL 480
Db 437 QNTGWSNANVKGTQIAYFGASLCSVDVDSNGSTDLVIGAPHYEOTRGGOVSCPL 496
481 PRGORARWQCDVAVLYGEOQFPGWGFAGALTVDVGVNGDKLTDVAGAPGEEDNRGAVYLF 540
Db 497 PRGORARWQCDVAVLYGEOQFPGWGFAGALTVDVGVNGDKLTDVAGAPGEEDNRGAVYLF 556
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
601 PVLRVKAIMBNPREVARNVFECDQVVKKEAGEVRVCLHVQKSTRDRLEGGQIQSVVT 660
Db 617 PVLRVKAIMBNPREVARNVFECDQVVKKEAGEVRVCLHVQKSTRDRLEGGQIQSVVT 676
661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
721 SLVGTPLSAFCNLRPVLAEDAQRLLFTALFPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFCNLRPVLAEDAQRLLFTALFPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 796
781 GPREFNVTYVRNDGEDSYRTQVTFPPFDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTYVRNDGEDSYRTQVTFPPFDLSYRKVSTLQNRQSRWRLACESASSTEV 856
```

```
QY 841 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKLVNNTSENNMPTNKTEP 900
Db 857 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKLVNNTSENNMPTNKTEP 916
QY 901 QLELPVKYAVYVWVTSYKLNFTASNTSRVWQHGYQVSNLQORSLPISLVFLVPV 960
Db 917 QLELPVKYAVYVWVTSYKLNFTASNTSRVWQHGYQVSNLQORSLPISLVFLVPV 976
QY 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVWNCSTAVCQRIQCIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVWNCSTAVCQRIQCIP 1036
QY 1021 PFGIQEFNATLKGNLSFDWYKTSNHLIIVSTAEILFNDVSTLLPQGFAPVRSQTEP 1080
Db 1037 PFGIQEFNATLKGNLSFDWYKTSNHLIIVSTAEILFNDVSTLLPQGFAPVRSQTEP 1096
QY 1081 KVEPFPVNPPLPIVGVSSVGGLLALLAITAALYKLGFFKQYKQKMMSEGGPPGABEQ 1137
Db 1097 KVEPFPVNPPLPIVGVSSVGGLLALLAITAALYKLGFFKQYKQKMMSEGGPPGABEQ 1153

RESULT 7
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Iedbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176

Query Match 99.5%; Score 5852; DB 14; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTPOENARFGQSVVQLQGSRVVVGAPQEIIVANQRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTPOENARFGQSVVQLQGSRVVVGAPQEIIVANQRGSLYQCDYSTGSCPEI 76
QY 61 RLOVPVEAVNMSLGLSLAATTPOLLACGPTVHTCSENTYVYKGLCFLFGSNLRQCPQK 120
Db 77 RLOVPVEAVNMSLGLSLAATTPOLLACGPTVHTCSENTYVYKGLCFLFGSNLRQCPQK 136
QY 121 PPEALRGCPQSDSDIAFLIDGSGSIIPHDPFRMKEWVSTVMEQLKXKTLFSLMQYSEEP 180
Db 137 PPEALRGCPQSDSDIAFLIDGSGSIIPHDPFRMKEFVSTVMEQLKXKTLFSLMQYSEEF 196
QY 181 RHIFTFKFPQNNPNRSLKBITOLLGHTHTATGLRKVVRELENTNGARKNAKILFLL 240
Db 197 RHIFTFKFPQNNPNRSLKBITOLLGHTHTATGLRKVVRELENTNGARKNAKILVVI 256
QY 241 TGEKFGDPLGYEDVIPELDREGVIRYVIGVDAPFRSEKRSQELNTVASKPPRDHVFQIN 300
Db 257 TGEKFGDPLGYEDVIPELDREGVIRYVIGVDAPFRSEKRSQELNTVASKPPRDHVFQIN 316
QY 301 NPEALKTIQNLREKIFAIETGTQGSSSSFEHMSQEGFSAITNSGELLSTVGSYDWAG 360
Db 317 NPEALKTIQNLREKIFAIETGTQGSSSSFEHMSQEGFSAITNSGELLSTVGSYDWAG 376
QY 361 GVFLYTSSEKSPFINWTVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSSEKSPFINWTVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTGGQSVUCPL 480
```

```
Db 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTGGQSVUCPL 496
QY 481 PRGQARWQCDVLYGQGGQWGFAGALTVLGDVNGDKLTDVAIGAPGEEDNRGAVTLF 540
Db 497 PRGQARWQCDVLYGQGGQWGFAGALTVLGDVNGDKLTDVAIGAPGEEDNRGAVTLF 556
QY 541 HGTSGSGISPSHSGRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAGQHVLILRSQ 600
Db 557 HGTSGSGISPSHSGRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAGQHVLILRSQ 616
QY 601 PVLRVKALMEFPNPREVARNVPECDQVYVKGKAGEVRVCLHVOKSTRDLREGQIQSVVT 660
Db 617 PVLRVKALMEFPNPREVARNVPECDQVYVKGKAGEVRVCLHVOKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPPVPIVRLNFP 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPPVPIVRLNFP 736
QY 721 SLVGTPLSAPGNLPRVLAEDAQRULTALPPPEKNCNDNI CODDLSITFFSMSLDCLVVG 780
Db 737 SLVGTPLSAPGNLPRVLAEDAQRULTALPPPEKNCNDNI CODDLSITFFSMSLDCLVVG 796
QY 781 GPREFNVTTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQHORSORSWRLACESASSTEV 840
Db 797 GPREFNVTTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQHORSORSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKLVNNTSENNMPTNKTEP 900
Db 857 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKLVNNTSENNMPTNKTEP 916
QY 901 QLELPVKYAVYVWVTSYKLNFTASNTSRVWQHGYQVSNLQORSLPISLVFLVPV 960
Db 917 QLELPVKYAVYVWVTSYKLNFTASNTSRVWQHGYQVSNLQORSLPISLVFLVPV 976
QY 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVWNCSTAVCQRIQCIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVWNCSTAVCQRIQCIP 1036
QY 1021 PFGIQEFNATLKGNLSFDWYKTSNHLIIVSTAEILFNDVSTLLPQGFAPVRSQTEP 1080
Db 1037 PFGIQEFNATLKGNLSFDWYKTSNHLIIVSTAEILFNDVSTLLPQGFAPVRSQTEP 1096
QY 1081 KVEPFPVNPPLPIVGVSSVGGLLALLAITAALYKLGFFKQYKQKMMSEGGPPGABEQ 1137
Db 1097 KVEPFPVNPPLPIVGVSSVGGLLALLAITAALYKLGFFKQYKQKMMSEGGPPGABEQ 1153

RESULT 8
US-09-902-481A-5
; Sequence 5, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimada, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RPT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-5
```

Query Match	99.4%;	Score 5851;	DB 10;	Length 1137;	
Best Local Similarity	99.3%;	Pred. No. 0;			
Matches 1129;	Conservative	5;	Mismatches	3;	Indels 0;
QY	1	FNLDTENAMTFOENARGFQSVVQVQGGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI	60		
DB	1	FNLDTENAMTFOENARGFQSVVQVQGGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI	60		
QY	61	RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQTCSNTYVVKGLCLFSGNLRRQPOK	120		
DB	61	RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQTCSNTYVVKGLCLFSGNLRRQPOK	120		
QY	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWSTVMEQKSKTSLFSLMOYSEF	180		
DB	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWSTVMEQKSKTSLFSLMOYSEF	180		
QY	181	RIHFTFKFQONNPNRSLIKPTQLLGRTHATGLRKVVRELFTNGARKNAFKILFL	240		
DB	181	RIHFTFKFQONNPNRSLIKPTQLLGRTHATGLRKVVRELFTNGARKNAFKILFL	240		
QY	241	TGGEKFGDPLGYEDVIPEDRGVIRYVIGVDAPRSEKSRQELNTVASKPPDRHVQIN	300		
DB	241	TGGEKFGDPLGYEDVIPEDRGVIRYVIGVDAPRSEKSRQELNTVASKPPDRHVQIN	300		
QY	301	NFEALKTIONLREKIFAIEGTQTCSSSFHEMSQBFSAIITNGPLLSWTGSDYDAG	360		
DB	301	NFEALKTIONLREKIFAIEGTQTCSSSFHEMSQBFSAIITNGPLLSWTGSDYDAG	360		
QY	361	GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAIILNRVQSLVLTGAPRYOHIGLVAMFR	420		
DB	361	GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAIILNRVQSLVLTGAPRYOHIGLVAMFR	420		
QY	421	QNTGWESNANVKGQIQIAYFGASICSDVDVDSNGSTDVLVIGAPHYYCTRGQGVSVCP	480		
DB	421	QNTGWESNANVKGQIQIAYFGASICSDVDVDSNGSTDVLVIGAPHYYCTRGQGVSVCP	480		
QY	481	PRGORARWQCDVLYGEOGPNGRFGAALTVDGVDNGDKLTVAICAPGEEDNRGAVYLF	540		
DB	481	PRGORARWQCDVLYGEOGPNGRFGAALTVDGVDNGDKLTVAICAPGEEDNRGAVYLF	540		
QY	1021	FFGIQEBFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTET	1080		
DB	1021	FFGIQEBFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTET	1080		
QY	1081	KVPEFVNPPLPLTVGSSVGGALLALITAAALYKLGFFKQYKDMSEGPPGABPQ	1137		
DB	1081	KVPEFVNPPLPLTVGSSVGGALLALITAAALYKLGFFKQYKDMSEGPPGABPQ	1137		
RESULT 9					
US-09-902-481A-6					
; Sequence 6, Application US/09902481A					
; Publication No. US2003005440A1					
; GENERAL INFORMATION:					
; APPLICANT: Springer, Timothy					
; APPLICANT: Shimaoka, Motomu					
; APPLICANT: Shifman, Julia					
; APPLICANT: Mayo, Stephen					
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY					
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK					
; CURRENT APPLICATION NUMBER: US/09/902.481A					
; CURRENT FILING DATE: 2001-07-09					
; PRIOR APPLICATION NUMBER: US 60/216,600					
; PRIOR FILING DATE: 2000-07-07					
; NUMBER OF SEQ ID NOS: 7					
; SOFTWARE: Patent in version 3.1					
; SEQ ID NO 6					
; LENGTH: 1137					
; TYPE: PRT					
; ORGANISM: Artificial sequence					
; FEATURE:					
; OTHER INFORMATION: synthetic					
US-09-902-481A-6					
Query Match 99.3%; Score 5845; DB 10; Length 1137;					
Best Local Similarity 99.0%; Pred. No. 0;					
Matches 1126; Conservative 8; Mismatches 3; Indels 0; Gaps 0;					
QY	1	FNLDTENAMTFOENARGFQSVVQVQGGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI	60		
DB	1	FNLDTENAMTFOENARGFQSVVQVQGGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI	60		
QY	61	RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQTCSNTYVVKGLCLFSGNLRRQPOK	120		
DB	61	RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQTCSNTYVVKGLCLFSGNLRRQPOK	120		
QY	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWSTVMEQKSKTSLFSLMOYSEF	180		
DB	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWSTVMEQKSKTSLFSLMOYSEF	180		
QY	181	RIHFTFKFQONNPNRSLIKPTQLLGRTHATGLRKVVRELFTNGARKNAFKILFL	240		
DB	181	RIHFTFKFQONNPNRSLIKPTQLLGRTHATGLRKVVRELFTNGARKNAFKILFL	240		
QY	241	TGGEKFGDPLGYEDVIPEDRGVIRYVIGVDAPRSEKSRQELNTVASKPPDRHVQIN	300		
DB	241	TGGEKFGDPLGYEDVIPEDRGVIRYVIGVDAPRSEKSRQELNTVASKPPDRHVQIN	300		
QY	301	NFEALKTIONLREKIFAIEGTQTCSSSFHEMSQBFSAIITNGPLLSWTGSDYDAG	360		
DB	301	NFEALKTIONLREKIFAIEGTQTCSSSFHEMSQBFSAIITNGPLLSWTGSDYDAG	360		
QY	361	GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAIILNRVQSLVLTGAPRYOHIGLVAMFR	420		
DB	361	GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAIILNRVQSLVLTGAPRYOHIGLVAMFR	420		
QY	421	QNTGWESNANVKGQIQIAYFGASICSDVDVDSNGSTDVLVIGAPHYYCTRGQGVSVCP	480		
DB	421	QNTGWESNANVKGQIQIAYFGASICSDVDVDSNGSTDVLVIGAPHYYCTRGQGVSVCP	480		
QY	481	PRGORARWQCDVLYGEOGPNGRFGAALTVDGVDNGDKLTVAICAPGEEDNRGAVYLF	540		
DB	481	PRGORARWQCDVLYGEOGPNGRFGAALTVDGVDNGDKLTVAICAPGEEDNRGAVYLF	540		


```
QY 541 HGTSGGSGISPSHSQRIAGSKSLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
DB 541 HGTSGGSGISPSHSQRIAGSKSLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
QY 601 PVLAVKALMEFNPREVAVNFECDQVVKGKEAGEVRVCLHVQKSTRDLRECGIQSVVT 660
DB 601 PVLAVKALMEFNPREVAVNFECDQVVKGKEAGEVRVCLHVQKSTRDLRECGIQSVVT 660
QY 661 YDLALDSGRPHSAVNETKNSRTRQTVGLGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720
DB 661 YDLALDSGRPHSAVNETKNSRTRQTVGLGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720
QY 721 SLVGTPLSAFNLRPVLAEDAORLFTALPFPFNKCGNDNICQDDLSITFSFMSLDCLVVG 780
DB 721 SLVGTPLSAFNLRPVLAEDAORLFTALPFPFNKCGNDNICQDDLSITFSFMSLDCLVVG 780
QY 781 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNSQSRWSRLACESASTEV 840
DB 781 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNSQSRWSRLACESASTEV 840
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKLVKANTVSENNMPRTNKTEF 900
DB 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKLVKANTVSENNMPRTNKTEF 900
QY 901 QLELPVKAYVMVTVSHGVSTKYLNETASENTSRVMQHOXOVSNLQORSLSPIVLRLNF 960
DB 901 QLELPVKAYVMVTVSHGVSTKYLNETASENTSRVMQHOXOVSNLQORSLSPIVLRLNF 960
QY 961 RLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDDIP 1020
DB 961 RLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDDIP 1020
QY 1021 FFGIOEEFNATLKGSLSPDNYIKTSHNHLLIVSTAEILLNDVSFTLLPGQGAFVRSQTE 1080
DB 1021 FFGIOEEFNATLKGSLSPDNYIKTSHNHLLIVSTAEILLNDVSFTLLPGQGAFVRSQTE 1080
QY 1081 KVEPPEVNPPLIIVGSSVGGLLILALITAALYKLGFFKQYKDMWSEGGPPGABPQ 1137
DB 1081 KVEPPEVNPPLIIVGSSVGGLLILALITAALYKLGFFKQYKDMWSEGGPPGABPQ 1137

RESULT 10
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimadzu, Motomu
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; FILE REFERENCE: DESIRED CONFORMATION AND METHODS FOR PRODUCING SAME
; CURRENT APPLICATION NUMBER: US/09/945,265
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,700
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-945-265-4

Query Match 99.2%; Score 5836.5; DB 9; Length 1152;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1127; Conservative 7; Mismatches 2; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPOBIVAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPOBIVAANQSGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSENTYVKGCLCFLFGSNLRQOPQ 120
```

```
DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSENTYVKGCLCFLFGSNLRQOPQ 136
QY 121 PPEARLARGCQPOSDIAFLIDGSGSIIPHDPRMKEMWVSTMEOQLKSKTLPGLMYSEEF 180
DB 137 PPEARLARGCQPOSDIAFLIDGSGSIIPHDPRMKEMWVSTMEOQLKSKTLPGLMYSEEF 196
QY 181 RIHFTHFKSQFQNNPNRSLIKPITQLLGRTHATGLRKVVVRELFINITNGARKNAFKILFLL 240
DB 197 RIHFTHFKSQFQNNPNRSLIKPITQLLGRTHATGLRKVVVRELFINITNGARKNAFKILVVI 256
QY 241 TDGKRFPGPLGYEDVTPELDREGVIRYVIGVGDASFBSKSRQELANTVASKPRDHVFQIN 300
DB 257 TDGKRFPGPLGYEDVTPELDREGVIRYVIGVGDASFBSKSRQELANTVASKPRDHVFQIN 316
QY 301 NFEALKTIQNLREKI PAIEGTQTGSSSSPHEMSQEGFSAAITNSGPLLSTVGSYDWA 360
DB 317 NFEALKTIQNLREKI PAIEGTQTGSSSSPHEMSQEGFSAAITNSGPLLSTVGSYDWA 376
QY 361 GVFPLTSKESKTFINMTRVDSMDDAYLYGAAAAIILNRVQSLVAGARYOHIGLVAMPR 420
DB 377 GVFPLTSKESKTFINMTRVDSMDDAYLYGAAAAIILNRVQSLVAGARYOHIGLVAMPR 436
QY 421 QNTGHWBSNANVKGQIIGAYFGASLCSVDVDSNGSTDVLVIGAPHYBQTRGQSVVCP 480
DB 437 QNTGHWBSNANVKGQIIGAYFGASLCSVDVDSNGSTDVLVIGAPHYBQTRGQSVVCP 496
QY 481 PRGQBARWQCDVAVLYGBOGPWGRFGAALTUVLDVNGDKLTDVAI GAPGEEDNRGA VILP 540
DB 497 PRGQBARWQCDVAVLYGBOGPWGRFGAALTUVLDVNGDKLTDVAI GAPGEEDNRGA VILP 555
QY 541 HGTSGGSGISPSHSQRIAGSKSLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
DB 556 HGTSGGSGISPSHSQRIAGSKSLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 615
QY 601 PVLAVKALMEFNPREVAVNFECDQVVKGKEAGEVRVCLHVQKSTRDLRECGIQSVVT 660
DB 616 PVLAVKALMEFNPREVAVNFECDQVVKGKEAGEVRVCLHVQKSTRDLRECGIQSVVT 675
QY 661 YDLALDSGRPHSAVNETKNSRTRQTVGLGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720
DB 676 YDLALDSGRPHSAVNETKNSRTRQTVGLGLTQTCETLKLQLPNCIEDPVSIVLRLNF 735
QY 721 SLVGTPLSAFNLRPVLAEDAORLFTALPFPFNKCGNDNICQDDLSITFSFMSLDCLVVG 780
DB 736 SLVGTPLSAFNLRPVLAEDAORLFTALPFPFNKCGNDNICQDDLSITFSFMSLDCLVVG 795
QY 781 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNSQSRWSRLACESASTEV 840
DB 796 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNSQSRWSRLACESASTEV 855
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKLVKANTVSENNMPRTNKTEF 900
DB 856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKLVKANTVSENNMPRTNKTEF 915
QY 901 QLELPVKAYVMVTVSHGVSTKYLNETASENTSRVMQHOXOVSNLQORSLSPIVLRLNF 960
DB 916 QLELPVKAYVMVTVSHGVSTKYLNETASENTSRVMQHOXOVSNLQORSLSPIVLRLNF 975
QY 961 RLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDDIP 1020
DB 976 RLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDDIP 1035
QY 1021 FFGIOEEFNATLKGSLSPDNYIKTSHNHLLIVSTAEILLNDVSFTLLPGQGAFVRSQTE 1080
DB 1036 FFGIOEEFNATLKGSLSPDNYIKTSHNHLLIVSTAEILLNDVSFTLLPGQGAFVRSQTE 1095
QY 1081 KVEPPEVNPPLIIVGSSVGGLLILALITAALYKLGFFKQYKDMWSEGGPPGABPQ 1137
DB 1096 KVEPPEVNPPLIIVGSSVGGLLILALITAALYKLGFFKQYKDMWSEGGPPGABPQ 1152
```

RESULT 11

US-10-116-275-204
; Sequence 204, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Eran Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrnes, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 204
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-204

Query Match 58.8%; Score 3459; DB 15; Length 1163;
Best Local Similarity 60.9%; Pred. No. 4.5e-310;
Matches 687; Conservative 142; Mismatches 294; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGFGQSVVOLGSRVWVGAPOEIVAAHQSGSLYQCYDSTGSCBPI 60
DB 20 FNLDTENAMTFOENARGFGQSVVOLGSRVWVGAPOEIVAAHQSGSLYQCYDSTGSCBPI 79

QY 61 RLOVPVEAVNMGLSLAATTSPPOLLACGPTVHOTCSENYVKGICELFGSNLRQPOK 120
DB 80 GLQVPEAVNMGLSLAATTSPPOLLACGPTVHOTCSENYVKGICELFGSNLRQPOK 137

QY 121 FPEALRGCPQEDSDIAFLDGSISIIIPHDFRMEKWSVTMEQLKSKTKLFSLMQYSEEF 180
DB 138 LPVSRQECRQEQDIIVFLDGSISISRNPFATWVNFVRAVISQFQSTQFSLMQFSNKF 197

QY 181 RHFTPEFQNNPNSLKIPIITOLGRTHATGLKRVRELNIINGARKNAFKILFLL 240
DB 198 QTHFTFEERFRTNPLSLASVHQLGQFTYATAIQVWVHRLPHASYGARRDATKILVI 257

QY 241 TDEKFGDPLGYDVIPELDREGVIRVIGVGDAPRSEKSRQELNTVWAKPRDHVQFQIN 300
DB 258 TDCKEGSDLDYKDVIPMDAAGIIRYAGVGLAFQNRNSWKLNDIAKSPQSEHIFKYE 317

QY 301 NFPAKTIQNLREKIPAIETGTGSSSPHEHMSOEGFSAATSNGLLSTVGSVDWAG 360
DB 318 DFDALKOIQNLKEKIPAIETGTGSSSPHEHMSOEGFSAATSNGLLSTVGSVDWAG 377

QY 361 GVFLYTSKESKSTINMTRVDSQNDVAVGAAAIILNRVQSLVGLAPRYOHIGLVAMER 420
DB 378 GAFLYPPNMSPTFINMSQENVMRDSYLGYSTELALWKGVQSLVGLAPRYOHIGLVAMER 437

QY 421 QNTGMENANVKTGTOIGYFGASCLSDVDVDSNGSTDVLIGAPHYYEQTGQSVQCP 480
DB 438 QVSRQEWKAEVGTQIGSYFGASCLSDVDVDSNGSTDVLIGAPHYYEQTGQSVQCP 497

QY 481 PRGQARWQCDVLYGEOGQWGRFGAALTIVLGVNCDKLTVAIGAPGEENRGAVYLF 540
DB 498 PRGWR-RWWCDVLYGEOGQWGRFGAALTIVLGVNCDKLTVAIGAPGEENRGAVYLF 556

QY 541 HGTSGSGISPSHSORLAGSKLSPRYFGOSLGGQDLTMDGLVLTVAQGHVLLLRQ 600
DB 557 HGVLGPSISPSHSORLAGSKLSPRYFGOSLGGQDLTMDGLVLTVAQGHVLLLRQ 616

QY 601 PVLVRKAIMENPREAVARNVFCNDQVVKGEAGEVRLVHVKSTRDLRBEQIOQSVTT 660
DB 617 PVLWVGVSQWQFPAEIPRSAFECEQVWSQTLVQSNICLYIDKRSKNLGSRLQSSVT 676

QY 661 YDLALDSGRPHSAVFNENKSTRTOTVGLTQTCETLKLQLPNCIEDPVSVILRLNF 720

DB 677 LDALDFCLSPRATFOETKNSLSRVRLGLKAKHCENFNLLPSCVEDSVTPITLRLNF 736
QY 721 SLVGTPLSAFCNLRPLAEDAQRLLTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 780
DB 737 TLVGKPLAFENLAPMLAALAQRYFTASLPPEKNOGADHICQDNLGISFPGKSLVVG 796
QY 781 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYKRVSTLQNRQSRWRLACESASSTEV 840
DB 797 SNLEINAEVWVWVNDGDSYRTQVTFPPPLDLSYKRVSTLQNRQSRWRLACESASSTEV 854
QY 841 SGALKSTSCSINHPIFFPENSFTNITFDVDSKASLGNKILLKANVTSENMMPTNKTEF 900
DB 855 SQGTWSTSCRINHPIFFPENSFTNITFDVDSKASLGNKILLKANVTSENMMPTNKTEF 914
QY 901 QLELPVKYAVVWVTVSHGVSTKYNFTAS-ENTSRVMOHOYOVSNLQORSPLISLFLVP 959
DB 915 QLELPVKYAVVWVTVSHGVSTKYNFTAS-ENTSRVMOHOYOVSNLQORSPLISLFLVP 974
QY 960 VRLNCTVLDWDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWVNGSIAVCORIQCDI 1019
DB 975 VELNOEAVWMDVEVSHPNQPSLRCSSEKIAFPASDFLAHIOKNPVLDCLSIACCLAFCDV 1034
QY 1020 PFFGQIEEFNATLKENLSPDWIKTSHNHLIVSTAILEFNDVSTFLPGQAFVRSOTE 1079
DB 1035 PSFSVQEEELDFTLKGNLSFGWVRQILQKKVSVSVAEITFDTSVYSQLPQGFQAFMRAQT 1094
QY 1080 TKVEPFEVNPFLPIVSGVGLLILLALITLALYKLGFEKQYKXDMMS 1128
DB 1095 TVLEKYKHNPFLPIVSGVGLLILLALITLALYKLGFEKQYKXDMMS 1143

RESULT 12
US-09-350-259-4
; Sequence 4, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-4

Query Match 58.4%; Score 3436; DB 9; Length 1163;
Best Local Similarity 60.6%; Pred. No. 6e-308;
Matches 684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGFGQSVVOLGSRVWVGAPOEIVAAHQSGSLYQCYDSTGSCBPI 60
DB 20 FNLDTENAMTFOENARGFGQSVVOLGSRVWVGAPOEIVAAHQSGSLYQCYDSTGSCBPI 79

QY 61 RLOVPVEAVNMGLSLAATTSPPOLLACGPTVHOTCSENYVKGICELFGSNLRQPOK 120
DB 80 GLQVPEAVNMGLSLAATTSPPOLLACGPTVHOTCSENYVKGICELFGSNLRQPOK 137

QY 121 FPEALRGCPQEDSDIAFLDGSISIIIPHDFRMEKWSVTMEQLKSKTKLFSLMQYSEEF 180

138 LPVSRQCPQEDIVFLIDGSGISSRNPFATMNFVRAVISQFQRPSTQFSLMQPSNKP 197
181 RHFTFKFQNNPNSLKIPIQLGLRTHATGLKAVVRELNFNITNGARKNAKILFL 240
198 QTHFTFEBFRTNPNLSLASVHQLQGFYTTATQNVVHRLPHASYGARRDAIKILVI 257
241 TDGEKFGDPLGYEDVPELDRGVIRYVIGVGDAPFRSEKSRQELNTVASKPRDHVFOIN 300
258 TDCKKSGDSLVDKVPIMADAAGIIRYAGVGLAFQNRNSWKELNDIAKPSQEHIFKVE 317
301 NPEALKTIQNLREKIPALIEGTOTGSSSFHEMSQEGFSAITNSGPNLLSTVGSVDWAG 360
318 DFDALDKIQNLKEKIPALIEGTETISSSFLEMAQEGFSAVTPDGPVLGAVGFTWSG 377
361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVLGAPRYQHIGLAVMPR 420
378 GAFLYPPNMSPTFINNSQENVNDRSDVYLGSTELALWKGVQSLVLGAPRYQHIGRAVIFI 437
421 QNTGHWESNANVKQTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSCVPL 480
438 QVSRQWRKARVIGTQIGSYFGASLCSVDVDTGSDTLVLIGAPHYEQTRGGQVSCVPL 497
481 PRGORARWOCDAVLYEGQGPWGRFGAALTVDGVNKGDLTDVAIGAPGEEDNRGAVYLF 540
498 PRGR-RWKCDAVLYEGQGPWGRFGAALTVDGVNKGDLTDVAIGAPGEEDNRGAVYLF 556
541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLGGQDLTDGVLDTVGAQGHVLLRQ 600
557 HGVLPSPISPSHQSRIAGSKLSPRLQYFGQSLGGQDLTDGVLDTVGAQGHVLLRTR 616
601 PVLRYKALMEFNPFEARNVPECNDQVYKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
617 PVLWVGVMQFIPASI PRSAFECRQVVSQETLVQSNICLYIDKRSKLLGSRDLQSSVT 676
661 YDIALDSGRPHSRAVFNETHKSTRQTOVLGLTQTCETILKLOLPNCIEDPSPVILRLNF 720
677 LDIALAPGRLSPRATFQETKRSLSRVRLGKACENFNLLPSCVEDSVIPIILRLNF 736

RESULT 13

US-09-891-943-4

; Sequence 4, Application US/09891943

; Publication No. US2003007728A1

; GENERAL INFORMATION:

; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US2003007728A1el Human 2
; FILE REFERENCE: 27666/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-891-943-4

Query Match 58.4%; Score 3436; DB 10; Length 1163;
Best Local Similarity 60.6%; Pred. No. 6e-308;
Matches 684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;

QY 1 FNLDTEANMTQENARGGQSVVQLQGSRVVVGAPQETVAANQSGSLYQCDYSTGSCPEI 60
DB 20 FNLDTEELTAFRVDSAGFGSDSVQVANSVVGAPQKIIAANQIGLQCGYSTGACPEI 79
QY 61 RIQVPEAVNMSLGLSLAATTPPOLLACGPTVHOTCSENTYVVKGLCFPLGSLNLRQOPK 120
DB 80 GLQVPEAVNMSLGLSLASTTSPSOLLACGPTVHHECGRNMYLTGLCLFLGPT--QLTOR 137
QY 121 PPEALRGCPQEDSDTAFLIDGSGSIIPDFPRMKWSTVMEOLKSKTSLSLMOYSSEF 180
DB 138 LPVSRQCPQEDIVFLIDGSGISSRNPFATMNFVRAVISQFQRPSTQFSLMQPSNKP 197
QY 181 RHFTFKFQNNPNSLKIPIQLGLRTHATGLKAVVRELNFNITNGARKNAKILFL 240
DB 198 QTHFTFEBFRTNPNLSLASVHQLQGFYTTATQNVVHRLPHASYGARRDAIKILVI 257
QY 241 TDGEKFGDPLGYEDVPELDRGVIRYVIGVGDAPFRSEKSRQELNTVASKPRDHVFOIN 300
DB 258 TDCKKSGDSLVDKVPIMADAAGIIRYAGVGLAFQNRNSWKELNDIAKPSQEHIFKVE 317
QY 301 NPEALKTIQNLREKIPALIEGTOTGSSSFHEMSQEGFSAITNSGPNLLSTVGSVDWAG 360
DB 318 DFDALDKIQNLKEKIPALIEGTETISSSFLEMAQEGFSAVTPDGPVLGAVGFTWSG 377
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVLGAPRYQHIGLAVMPR 420
DB 378 GAFLYPPNMSPTFINNSQENVNDRSDVYLGSTELALWKGVQSLVLGAPRYQHIGRAVIFI 437
QY 421 QNTGHWESNANVKQTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSCVPL 480
DB 438 QVSRQWRKARVIGTQIGSYFGASLCSVDVDTGSDTLVLIGAPHYEQTRGGQVSCVPL 497
QY 481 PRGORARWOCDAVLYEGQGPWGRFGAALTVDGVNKGDLTDVAIGAPGEEDNRGAVYLF 540
DB 498 PRGR-RWKCDAVLYEGQGPWGRFGAALTVDGVNKGDLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLGGQDLTDGVLDTVGAQGHVLLRQ 600
DB 557 HGVLPSPISPSHQSRIAGSKLSPRLQYFGQSLGGQDLTDGVLDTVGAQGHVLLRTR 616
QY 601 PVLRYKALMEFNPFEARNVPECNDQVYKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLWVGVMQFIPASI PRSAFECRQVVSQETLVQSNICLYIDKRSKLLGSRDLQSSVT 676
QY 661 YDIALDSGRPHSRAVFNETHKSTRQTOVLGLTQTCETILKLOLPNCIEDPSPVILRLNF 720
DB 677 LDIALAPGRLSPRATFQETKRSLSRVRLGKACENFNLLPSCVEDSVIPIILRLNF 736

QY 721 SLVGTPLSAFNLPRVLAEDAORLFTALPPFPKNCNDNICODLSITPFSMSLDCLVVG 780
DB 737 TLVGPPLAFLRNLPRMLAARQYTTASLPFRNCGADHICODNLGIFSPPLKSLVVG 796
QY 781 GPRFNVTVVRNDGDSYRTQVTFPPFLDLRYKRVSTLQNRQSRQSWRLACESASSTEV 840
DB 797 SNLEKAEVWVWVNDGDSYGTITFTSHDAGLSYRVVAEQKQQLRSLHLTC--CSAPVG 854
QY 841 SCALKSTSCSINHPIPPENSEVTENTDVSQKSLGNKLLKANVTSENMPRTNKTEP 900
DB 855 SGTWSTSCRINHLIFRGAQITFLATPDVSKAVGLDRLLIANVSSNNIPRTSKTIP 914
QY 901 QLELPVKAVYVTVSHGVSTKYLAFAS-ENTSRVMOHQYVSNLQORSPLISLAVFLVP 959
DB 915 QLELPVKAVYVTVSHGVSTKYLAFAS-ENTSRVMOHQYVSNLQORSPLISLAVFLVP 974
QY 960 VRLNQTVWDRPQVTFSENLSTCTHKLPSHSDPLAELRKAPVNVCSIAVCORIOCDI 1019
DB 975 VELNQBAVMDVEVHPQNSURCSSEKTAIPASDPLAHQKNPVLDCSIAQCLPRCDV 1034
QY 1020 PFFGIQEEFNATLKNLSPDWIKTSHNHLIVSTAELFNDVSFTLLPGQCAFVRSQTE 1079
DB 1035 PFSVQBELDPTLKNLSPDWIKTSHNHLIVSTAELFNDVSFTLLPGQCAFVRSQTE 1094
QY 1080 TKVPEFVNPPLPLVGVSSVGLLILALITAAALYKLGFPFKRQYKQWMBE 1128
DB 1095 TVLEKYVHNPIPLVGVSSVGLLILALITAAALYKLGFPFKRQYKQWMBE 1143

RESULT 14

US-09-350-259-2

; Sequence 2, Application US/09350259

; Patent No. US20020062008A1

; GENERAL INFORMATION:

; APPLICANT: Gallatin, Michael W.

; TITLE OF INVENTION: No. US20020062008A1el Human 2

; FILE REFERENCE: 27866/35004

; CURRENT APPLICATION NUMBER: US/09/350,259

; CURRENT FILING DATE: 1999-07-08

; EARLIER APPLICATION NUMBER: 09/193,043

; EARLIER FILING DATE: 1998-11-16

; EARLIER APPLICATION NUMBER: 08/173,497

; EARLIER FILING DATE: 1993-12-23

; EARLIER APPLICATION NUMBER: 08/286,889

; EARLIER FILING DATE: 1994-08-05

; EARLIER APPLICATION NUMBER: 08/362,652

; EARLIER FILING DATE: 1994-12-21

; EARLIER APPLICATION NUMBER: 08/943,363

; EARLIER FILING DATE: 1997-10-03

; NUMBER OF SEQ ID NOS: 114

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 1161

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-350-259-2

Query Match 57.8%; Score 3401; DB 9; Length 1161;

Best Local Similarity 59.3%; Pred. No. 1.1e-304;

Matches 669; Conservative 167; Mismatches 285; Indels 9; Gaps 6;

QY 1 FNLDTENAMTFOENARFGQSVVQLQSGRVVVGAPQEIYVAANRQSLYQCDYSTGSCPEI 60

DB 17 FNLDVEEPTIQEDAGGFGQSVVQVGGSRVVGAPLEVVAAVQGRVYDCAATGMCQPI 76

QY 61 RLQVPEAVNWSLGLSLAATTPPOLLAGCTVHTQCSNTYVGLCLPFGSNLRQOPQK 120

DB 77 FLHIREAVNWSLGLTAASTNGSLRACGPTLHRCVGENSYKSGSCILLGSRW-ETIQT 135

QY 121 PPEALRGCPQEDSDLAFLIDGSGSIIPHDFFRMKEWSTVMEOLKSKTLFSLAQYSEEF 180

DB 121 PPEALRGCPQEDSDLAFLIDGSGSIIPHDFFRMKEWSTVMEOLKSKTLFSLAQYSEEF 180

RESULT 15

US-09-891-943-2

; Sequence 2, Application US/09891943

; Publication No. US2003007278A1

; GENERAL INFORMATION:

; APPLICANT: Gallatin, Michael W.

DB 136 VPDATPECPHQEMDIVFLIDGSGSIDQNDPROMKGFVQAVMGQFEGTDTLTFALMQSNLL 195
QY 181 RIHTTPEFQONPNRSLIKPIITOLLGRTHTATGLRVKVRLEFNITWGARNKAKILLPLL 240
DB 196 KIHTTFTQFRTSPQSSQSLVDPIVOLKGLTFTATGILLTVVTLFHHKNGARKSARKILLIVI 255
QY 241 TDGKFGDPLGYEDVPELDREGVIRYVIGVDGAFRSEKSRQELMTVASKEPRDRHVFQIN 300
DB 256 TDGKYNDPLGYSDVIPAQKAGIIRYVIGVDGAFRSEKSRQELMTVASKEPRDRHVFQIN 315
QY 301 NFEALKTIQOLRKKIPIAIBCTQTGSSSSPEHEMSQEGFSAATNSNGPLLTSTVCSYDAG 360
DB 316 NFAALGSIQOLRKKIPIAIBCTQTGSSSSPEHEMSQEGFSAATNSNGPLLTSTVCSYDAG 375
QY 361 GVFLYTSKEKSTFNTMTTRVDSMDNDAYLGYAAAILLENRVQSLVGLGAPRYOHIGLVAMFR 420
DB 376 GAFLYPPNMSPTPINMSQENVDMEDSYLGYSTELALWQVQNLVGLGAPRYOHTKAVIFT 435
QY 421 QNTGWESNANVKTQIYGAFGASLCSVDVDSNGSTDLVIGAPHYTEOTRGGQSVCP 480
DB 436 QVSRQWRKKAETVQTQISYFASLCSVDVDSNGSTDLVIGAPHYTEOTRGGQSVCP 495
QY 481 PRGORARWQCDVLYGBOGOWPGRFGAALTIVLGVNDGDKLTDVAIGAPGEDNNGAVYLF 540
DB 496 PRGORVQWQCDVLYGBOGOWPGRFGAALTIVLGVNDGDKLTDVAIGAPGEDNNGAVYLF 555
QY 541 HGTSGSGISPSHSGRIAGSKLSPRLPYGQSLSGQDLTMDGLVDLTVGAGHVLRLRSQ 600
DB 556 HGASEGSGISPSHSGRIAGSKLSPRLPYGQSLSGQDLTMDGLVDLTVGAGHVLRLRSQ 615
QY 601 PVLAVKALMEPNPREVARNVTECDNDVVKGEAGEVVKLVQKSTRDLRLEGGOIQSVVT 660
DB 616 PVLKGVAMRFPSPVEVAKAVTRCWEKPSALEAGDATVCLTIQKSLDQL--GDIQSVR 673
QY 661 YDLALDSGRPHSRAVFNETKNSTRRTQVLGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720
DB 674 FDLALDPGLTSTRAI FNETKNPTLTKLGLTHCETLKLKLLPDCVEDVVSPIILNLP 733
QY 721 SLVGTPLSAFNLPRVLAEDAORLFTALPPFPKNCNDNICODLSITPFSMSLDCLVVG 780
DB 734 SLVREPIPSQNLPRVLAAGSQDLFTASLPFRNCGADHICODNLGIFSPPLKSLVVG 793
QY 781 GPRFNVTVVRNDGDSYRTQVTFPPFLDLRYKRVSTLQNRQSRQSWRLACESASSTEV 840
DB 794 SSELNVITVWVWVNDGDSYGTITFTSHDAGLSYRVVAEQKQQLRSLHLTC--CSAPVG 854
QY 841 SCALKSTSCSINHPIPPENSEVTENTDVSQKSLGNKLLKANVTSENMPRTNKTEP 900
DB 855 SGTWSTSCRINHLIFRGAQITFLATPDVSKAVGLDRLLIANVSSNNIPRTSKTIP 914
QY 901 QLELPVKAVYVTVSHGVSTKYLAFAS-ENTSRVMOHQYVSNLQORSPLISLAVFLVP 959
DB 912 QLELPVKAVYVTVSHGVSTKYLAFAS-ENTSRVMOHQYVSNLQORSPLISLAVFLVP 971
QY 960 VRLNQTVWDRPQVTFSENLSTCTHKLPSHSDPLAELRKAPVNVCSIAVCORIOCDI 1019
DB 972 VLLNGVAVMDVWVWVNDGDSYGTITFTSHDAGLSYRVVAEQKQQLRSLHLTC--CSAPVG 1029
QY 1020 PFFGIQEEFNATLKNLSPDWIKTSHNHLIVSTAELFNDVSFTLLPGQCAFVRSQTE 1079
DB 1030 PFSVQBELDPTLKNLSPDWIKTSHNHLIVSTAELFNDVSFTLLPGQCAFVRSQTE 1089
QY 1080 TKVPEFVNPPLPLVGVSSVGLLILALITAAALYKLGFPFKRQYKQWMBE 1128
DB 1090 TVLEKYVHNPIPLVGVSSVGLLILALITAAALYKLGFPFKRQYKQWMBE 1138

RESULT 15

US-09-891-943-2

; Sequence 2, Application US/09891943

; Publication No. US2003007278A1

; GENERAL INFORMATION:

; APPLICANT: Gallatin, Michael W.

721 SLVGTPLSFGNLPVLAEDAORLFTALFPFKKNCNDNICODDLSTFFSMISLCLVVG 780
 734 SLVREPIFSPQNLAPVLAAGSODLFTASLPFKKNCQDGLCEGLGVTLSFGSLQTLTVG 793
 781 GPREFNVTVVNDGDSYRTQVTFPPFLDI SVKVKSTLQKRSQRWRWLACESASSTEV 840
 794 SSLELNVITVWNAAGDSYGTVVSLLYPAGLSHRRVSGAQKQPHQSALELACETV-PTED 852
 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEP 900
 853 EG-LRSSRCSVNHPIFHEGSGNGTIFVTFDVSKATLGRMLMRASASSENKASSSKATP 911
 901 QLELPVKYAVVMVYVTSHGVSSTKYNLF-TASENTSRRVMQHOYQVSNLQORSIFISLVLVP 959
 912 QLELPVKYAVVMVYVTSHGVSSTKYNLF-TASENTSRRVMQHOYQVSNLQORSIFISLVLVP 971
 960 VRLNQTWVDRPQVTFSEHLSSTCHTKERLP SHSDDFLABLKAPVNCSTAVCQRIQCDI 1019
 972 VLLNGVAVWVWMEAPSQSL--PCVSEKPPQHSDFLTQISRSFMDLDCSIADCLQFRCDV 1029
 1020 PFFGIQSEFNATLKENLSFDWYIKTSHNHLIVSTABILFNDISVFTLLPGQGAFFVRSQTE 1079
 1030 PFSVQSELDFTLKENLSFGWVRETLQKVLVWSVAITFDTSVYSQLPQGEAFWEAQME 1089
 1080 TKVEPFEPVNPPLIVGSSVGLLLALITAAALYKLGFFKQYKDMMSSE 1128
 1090 MVLBEDEVYNAIPIIMGSSVGCALLLALITATLYKLGFFKQYKDMMSSE 1138

Search completed: June 7, 2004, 17:38:47
 Job time : 41.8905 secs

APPLICANT: Van der Vieren, Monica
 TITLE OF INVENTION: No. US2003007728A1el Human 2
 CURRENT PILING DATE: 2001-06-26
 CURRENT PILING DATE: 2001-06-26
 PRIOR PILING DATE: 1998-11-16
 PRIOR PILING DATE: 1998-11-16
 PRIOR PILING DATE: 1994-08-05
 PRIOR PILING DATE: 1994-08-05
 PRIOR PILING DATE: 1994-12-21
 PRIOR PILING DATE: 1994-12-21
 PRIOR PILING DATE: 1997-10-03
 PRIOR PILING DATE: 1997-10-03
 NUMBER OF SEQ ID NOS: 114
 SOFTWARE: Patent in Ver. 2.0
 LENGTH: 1161
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-891-943-2

Query Match 57.8%; Score 3401; DB 10; Length 1161;
 Best Local Similarity 59.3%; Pred. No. 1.1e-304;
 Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;

1 ENLDTENAMTFORNARFGOSVVLQGSRVVVGAPDQEIIVAAORGLSYOCYSTGSCBPI 60
 17 FNLDDVEPTIFQDAGGQSVVQFQGSRLVVGAPLEVVAAQTGELDYCAATGKQPI 76
 61 RLQVPVAVNMSLGLSIAATTSPQLACGTPVHQCSTENTYVKGCLFPGSNLQKQPK 120
 77 PLHIREAVNMSLGLTAASTNGSRLLAGCGTLHRCVCGNSYSGKSCLLGSRW-EIIQT 135
 121 FPEALRGCCPEDSDIAFLDGGSIIPHPFRMKWVSTVMEQLKSKTLPISLMOYSEEF 180
 136 VPATPECHQEMDIIVFLDGGSIDQDNFQNMKGQVQAVMGQFEGDTLTFALMQYSNLL 195
 181 RIHTEFKFQNNPRSLIKPIQLLGRTHRTATGLKRVVRELFTNIGARKNAFKILFL 240
 196 KIHFTTQRTSPSQSLVDPIVOLKGLTFTATGILTVTQTFHKGARKSAKILIVI 255
 241 TDGEKGDPLGHDVTELDREGVIRVYGVGDAPFSEKROELNIVASKPDRDHVQIN 300
 256 TDGQKYKDPLEYSVDVPOAKAGIIRIYAGVGHAFQPTARQELNTISSAPPQDHFVRVD 315
 301 NFALKTIQNLREKIFALEGTGTGSSSFHEMSQEGFSAATISNGPLLLSTVGSYDMAG 360
 316 NFAALGSIQKLOEKIYAVESTQSRASSSFQHEMSQEGFSTALTMGDLFLGAVGSFWSG 375
 361 GVPLTYSKEKSTPINTRVDSMDNDAYLGAAAIILBNRVOSLVIGAPRYOHIGLVAMFR 420
 376 GAFLYPPNMSPTFNKSNQENVDNRDYLGYSTELALWQVNLVLGAPRYOHTGKAVIF 435
 421 QNTGMWESNANVKTGTGAYFGASLCSVDVDSNGSTDLVLIAGAPHYETQTRGGQSVCP 480
 436 QVSRQWRKRAEVTGTQISYFGASLCSVDVDSNGSTDLVLIAGAPHYETQTRGGQSVCP 495
 481 PROGRARWQCDVLYGEGQGPWRGFGAALTVLGVNGDKLTDVAIGAPGEENRGAVYLF 540
 496 PRGQVQWQCDVAVLRGEQHPWGFAGALTVLGVNEDKLIDVAIGAPGEENRGAVYLF 555
 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLTVGAQGHVLLRSQ 600
 556 HGASESGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLTVGAQGHVLLRSQ 615
 601 PVLRVKAIMEFNPREVARNVFCNDQVVKGEAGEVVRVCLHVQKSTRDLREGQIOSVVT 660
 616 PVLKVGVAKEFSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673
 661 YDLALDSGRHSAVNETKSTNRTQVTLGTCTETKQLPNCIEDPVSPIVLRNLF 720
 674 FDLALDPGRLTSAIFNETKNPTLTKRKTGLGHCETLKLILLPDCVEDVWSPFIILHLP 733

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:07:20 ; Search time 18.3484 Seconds

(without alignments)
3199.127 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTQENARGFGQ.....FKROYKDWMBEGGPPGAEPQ 1137

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents 2A.*

1: /cgn2_6/ptodata/2/1aa/5A-COMB.pap:*

2: /cgn2_6/ptodata/2/1aa/5B-COMB.pap:*

3: /cgn2_6/ptodata/2/1aa/6A-COMB.pap:*

4: /cgn2_6/ptodata/2/1aa/6B-COMB.pap:*

5: /cgn2_6/ptodata/2/1aa/PTUS-COMB.pap:*

6: /cgn2_6/ptodata/2/1aa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5852	99.5	1153	1	US-08-173-497-3 Sequence 3, Appli
2	5852	99.5	1153	1	US-08-286-889-3 Sequence 3, Appli
3	5852	99.5	1153	1	US-08-485-618-3 Sequence 3, Appli
4	5852	99.5	1153	1	US-08-362-652-2 Sequence 3, Appli
5	5852	99.5	1153	2	US-08-605-672-3 Sequence 3, Appli
6	5852	99.5	1153	2	US-08-482-293A-3 Sequence 3, Appli
7	5852	99.5	1153	2	US-08-943-363-3 Sequence 3, Appli
8	5852	99.5	1153	3	US-09-193-043-3 Sequence 3, Appli
9	5852	99.5	1153	4	US-09-688-307A-3 Sequence 3, Appli
10	5852	99.5	1153	4	US-09-350-259-3 Sequence 3, Appli
11	5821.5	98.9	1152	2	US-08-476-062A-43 Sequence 43, Appli
12	5821.5	98.9	1152	2	US-08-476-062A-43 Sequence 43, Appli
13	5821.5	98.9	1152	6	US-08-476-062A-43 Patent No. 5424399
14	3459	58.8	1163	2	US-08-476-062A-44 Sequence 44, Appli
15	3459	58.8	1163	5	US-08-476-062A-44 Sequence 44, Appli
16	3436	58.4	1163	1	US-08-173-497-4 Sequence 4, Appli
17	3436	58.4	1163	1	US-08-286-889-4 Sequence 4, Appli
18	3436	58.4	1163	1	US-08-485-618-4 Sequence 4, Appli
19	3436	58.4	1163	1	US-08-362-652-4 Sequence 4, Appli
20	3436	58.4	1163	2	US-08-605-672-4 Sequence 4, Appli
21	3436	58.4	1163	2	US-08-482-293A-4 Sequence 4, Appli
22	3436	58.4	1163	2	US-08-943-363-4 Sequence 4, Appli
23	3436	58.4	1163	3	US-09-193-043-4 Sequence 4, Appli
24	3436	58.4	1163	4	US-09-688-307A-4 Sequence 4, Appli
25	3436	58.4	1163	4	US-09-350-259-4 Sequence 4, Appli
26	3401	57.8	1161	1	US-08-173-497-2 Sequence 2, Appli
27	3401	57.8	1161	1	US-08-286-889-2 Sequence 2, Appli

28	3401	57.8	1161	1	US-08-485-618-2	Sequence 2, Appli
29	3401	57.8	1161	1	US-08-362-652-2	Sequence 2, Appli
30	3401	57.8	1161	2	US-08-605-672-2	Sequence 2, Appli
31	3401	57.8	1161	2	US-08-482-293A-2	Sequence 2, Appli
32	3401	57.8	1161	2	US-08-943-363-2	Sequence 2, Appli
33	3401	57.8	1161	3	US-09-193-043-2	Sequence 2, Appli
34	3401	57.8	1161	4	US-09-688-307A-2	Sequence 2, Appli
35	3401	57.8	1161	4	US-09-350-259-2	Sequence 2, Appli
36	3385.5	57.5	1161	1	US-08-485-618-99	Sequence 99, Appli
37	3385.5	57.5	1161	2	US-08-605-672-99	Sequence 99, Appli
38	3385.5	57.5	1161	2	US-08-482-293A-99	Sequence 99, Appli
39	3385.5	57.5	1161	2	US-08-943-363-99	Sequence 99, Appli
40	3385.5	57.5	1161	3	US-09-193-043-99	Sequence 99, Appli
41	3385.5	57.5	1161	4	US-09-688-307A-99	Sequence 99, Appli
42	3385.5	57.5	1161	4	US-09-350-259-99	Sequence 99, Appli
43	3224.5	54.8	1161	3	US-09-193-043-55	Sequence 55, Appli
44	3224.5	54.8	1161	4	US-09-688-307A-55	Sequence 55, Appli
45	3224.5	54.8	1161	4	US-09-350-259-55	Sequence 55, Appli

ALIGNMENTS

RESULT 1
US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; TITLE OF INVENTION: No. 5437958 Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60608-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08173497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTQENARGFGQSVVQLQSGRVVVGAPQEIIVANQSGSLVQCYSTGSCBPI 60
DB 17 FNLDTENAMTQENARGFGQSVVQLQSGRVVVGAPQEIIVANQSGSLVQCYSTGSCBPI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQPOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQPOK 136
QY 121 FPALRGCPQEDSDIAFLIDGSGIIPHPFRMKWVSTVMEOLKSKTLPFLMOYSEEP 180
DB 137 FPALRGCPQEDSDIAFLIDGSGIIPHPFRMKWVSTVMEOLKSKTLPFLMOYSEEP 196
QY 181 RHFTFKFQNNPNRSLKPIITQLGRTHATGLRKVVRELFTNGARKNAKILFL 240
DB 197 RHFTFKFQNNPNRSLKPIITQLGRTHATGLRKVVRELFTNGARKNAKILFL 256
QY 241 TDCEKFGDPLGYEDVPEADREGVIRVVGDPAPRSEKSRQELNTVASKPRDHVFQIN 300
DB 257 TDCEKFGDPLGYEDVPEADREGVIRVVGDPAPRSEKSRQELNTVASKPRDHVFQIN 316
QY 301 NFPAKTIQNLREKIPAEIGTGTGSSSPHEHMSQEGFSAATISNGPLSTVGSYDWA 360
DB 317 NFPAKTIQNLREKIPAEIGTGTGSSSPHEHMSQEGFSAATISNGPLSTVGSYDWA 376
QY 361 GVFLYTSKSKSTINTRVDSNDAYLGYAAIILNRVQSLVLCAPRVOHIGLVAMER 420
DB 377 GVFLYTSKSKSTINTRVDSNDAYLGYAAIILNRVQSLVLCAPRVOHIGLVAMER 436
QY 421 QNTGWESNANVKGTOIGAFYFASLCSVDVDSNGSTDVLIGAPHYETQTRGQVSVCP 480
DB 437 QNTGWESNANVKGTOIGAFYFASLCSVDVDSNGSTDVLIGAPHYETQTRGQVSVCP 496
QY 481 PRGRARWQCDVLYGEOGPWGRFGAALTVLGVNGDKLTVAI GARGEDNRGAVYLF 540
DB 497 PRGRARWQCDVLYGEOGPWGRFGAALTVLGVNGDKLTVAI GARGEDNRGAVYLF 556
QY 541 HGTSGSGISGSHSQRISAGSKSLPRLOYFGQSLGGQDLTMDGLVLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISGSHSQRISAGSKSLPRLOYFGQSLGGQDLTMDGLVLTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIMFPNPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLRREGIOQSVT 660
DB 617 PVLRVKAIMFPNPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLRREGIOQSVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRQTVGLTQCTETLKLQPCNIEDPVSIVLRLNF 720
DB 677 YDLALDSGRPHSRVAFNETKSTRQTVGLTQCTETLKLQPCNIEDPVSIVLRLNF 736
QY 721 SLVGTPLSAGNLRPLVLAEDAQLFTALPFPBKNGNDNICDDLSITFSPMSLDCILV 780
DB 737 SLVGTPLSAGNLRPLVLAEDAQLFTALPFPBKNGNDNICDDLSITFSPMSLDCILV 796
QY 781 GPREFNVTVRNDGEDSYTQVTFPPDLVSRKYSTLONORSQSRWLACASSTEV 840
DB 797 GPREFNVTVRNDGEDSYTQVTFPPDLVSRKYSTLONORSQSRWLACASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEP 900
DB 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEP 916
QY 901 QLELPVYAVYVMTSHGVSTKYLNTASNTSRVMOHOYQVNLGORSIPISLVELVPV 960
DB 917 QLELPVYAVYVMTSHGVSTKYLNTASNTSRVMOHOYQVNLGORSIPISLVELVPV 976
QY 961 RLNTVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PNVNCSIAVCQRIQDIP 1020
DB 977 RLNTVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PNVNCSIAVCQRIQDIP 1036
QY 1021 PFGIQEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSTVFTLLPGQAFVRSQTET 1080
DB 1037 PFGIQEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSTVFTLLPGQAFVRSQTET 1096
QY 1081 KVPEPEVPNPLIIVGSSVGLLILALITAALYKLGFFKQYKDMWSEGGPPGAEPQ 1137
DB 1097 KVPEPEVPNPLIIVGSSVGLLILALITAALYKLGFFKQYKDMWSEGGPPGAEPQ 1153

RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286.889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-286-889-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQETVAANQSGSLYQCDYSTGSCRP 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQETVAANQSGSLYQCDYSTGSCRP 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQPOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQPOK 136
QY 121 FPALRGCPQEDSDIAFLIDGSGIIPHPFRMKWVSTVMEOLKSKTLPFLMOYSEEP 180
DB 137 FPALRGCPQEDSDIAFLIDGSGIIPHPFRMKWVSTVMEOLKSKTLPFLMOYSEEP 196
QY 181 RHFTFKFQNNPNRSLKPIITQLGRTHATGLRKVVRELFTNGARKNAKILFL 240
DB 197 RHFTFKFQNNPNRSLKPIITQLGRTHATGLRKVVRELFTNGARKNAKILFL 256
QY 241 TDCEKFGDPLGYEDVPEADREGVIRVVGDPAPRSEKSRQELNTVASKPRDHVFQIN 300
DB 257 TDCEKFGDPLGYEDVPEADREGVIRVVGDPAPRSEKSRQELNTVASKPRDHVFQIN 316
QY 301 NFPAKTIQNLREKIPAEIGTGTGSSSPHEHMSQEGFSAATISNGPLSTVGSYDWA 360
DB 317 NFPAKTIQNLREKIPAEIGTGTGSSSPHEHMSQEGFSAATISNGPLSTVGSYDWA 376

Db 557 HGTSGGSPSHSRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDTLTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRREGQIQSVVT 660
Db 617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQLPNCIEDPVSIVLRNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQLPNCIEDPVSIVLRNF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFKKNCNDNIQDDLSITFSFMSLDCLVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFKKNCNDNIQDDLSITFSFMSLDCLVG 796
Qy 781 GPREFNVTVTRDGEDSYRTQVTFPPPLDLSYRKVSTLQNSQSRWRLACSSASTEV 840
Db 797 GPREFNVTVTRDGEDSYRTQVTFPPPLDLSYRKVSTLQNSQSRWRLACSSASTEV 856
Qy 841 SGALKSTSCSINPIIPFENSEVFNITFDVDSKASLGNKLLKXANTVSENNMPTNKTEF 900
Db 857 SGALKSTSCSINPIIPFENSEVFNITFDVDSKASLGNKLLKXANTVSENNMPTNKTEF 916
Qy 901 QLELPVKYAVVMVTSHGVTCKYLINFASNTSRVMQHQYQVSNLQORSIPISLVLVPV 960
Db 917 QLELPVKYAVVMVTSHGVTCKYLINFASNTSRVMQHQYQVSNLQORSIPISLVLVPV 976
Qy 961 RLNQTVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVMNCSTAVCORIOCDIP 1020
Db 977 RLNQTVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVMNCSTAVCORIOCDIP 1036
Qy 1021 PFGIQEAFNATLGNLSFDWYIKTSHNHLIVSTAEILFNDSTVFTLLPGQGA FVRSQTET 1080
Db 1037 PFGIQEAFNATLGNLSFDWYIKTSHNHLIVSTAEILFNDSTVFTLLPGQGA FVRSQTET 1096
Qy 1081 KVBEPPEVNPPLIVSSVGGILLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1137
Db 1097 KVBEPPEVNPPLIVSSVGGILLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1153

RESULT 4
US-08-362-652-3
; Sequence 3, Application US/08362652
; Patent No. 5786850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 278666/32391
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-362-652-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVANQORSLQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVANQORSLQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFPGNLNQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFPGNLNQOPQK 136
Qy 121 FPEALRGCPQEDSDTAFIDGSGSIIPHDPRMKEMVSTVMEQLKSKTFLSLMOYSBEF 180
Db 137 FPEALRGCPQEDSDTAFIDGSGSIIPHDPRMKGFVSTVMEQLKSKTFLSLMOYSBEF 196
Qy 181 RIHFTFKGQFQNNPNRSLIKPITQLLGRTHATGKRVVRELFNITNGARKNAKILVLI 240
Db 197 RIHFTFKGQFQNNPNRSLIKPITQLLGRTHATGKRVVRELFNITNGARKNAKILVLI 256
Qy 241 TDGEKGGPLGVEDVIPDLREGVIRYVGVGDAPRSKROELNNTVASKPRDHVQIN 300
Db 257 TDGEKGGPLGVEDVIPDLREGVIRYVGVGDAPRSKROELNNTVASKPRDHVQIN 316
Qy 301 NFEALKTIQNLREKIPAIETGQTGSSSFEHMSQEGFSAAITSNGPLLTSGVSDWAG 360
Db 317 NFEALKTIQNLREKIPAIETGQTGSSSFEHMSQEGFSAAITSNGPLLTSGVSDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIIILNRVQSLVGLGAPRYOHIGLVAMFR 436
Qy 421 QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEBQTRGGQVSVCP 480
Db 437 QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEBQTRGGQVSVCP 496
Qy 481 PRGQARWQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPSEENRGAVYLF 540
Db 497 PRGQARWQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPSEENRGAVYLF 556
Qy 541 HGTSGGSPSHSRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDTLTVGAQGHVLLRSQ 600
Db 557 HGTSGGSPSHSRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDTLTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRREGQIQSVVT 660
Db 617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQLPNCIEDPVSIVLRNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQLPNCIEDPVSIVLRNF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFKKNCNDNIQDDLSITFSFMSLDCLVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFKKNCNDNIQDDLSITFSFMSLDCLVG 796
Qy 781 GPREFNVTVTRDGEDSYRTQVTFPPPLDLSYRKVSTLQNSQSRWRLACSSASTEV 840

Db 797 GPREFNVTVVRNDGDSYKQWTFPPDLDSYRKVSTIQNRSQSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTENITPDVDSKASLGKLLKXAVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTENITPDVDSKASLGKLLKXAVTSENMPRTNKTEF 916
Qy 901 QLELPVKYAVVMVTVSHGVSTKYLANFTASENTSRVMQHOYQVSNLQORSPLISLVLFPV 960
Db 917 QLELPVKYAVVMVTVSHGVSTKYLANFTASENTSRVMQHOYQVSNLQORSPLISLVLFPV 976
Qy 961 RLNOTVIMDRPQVTFSENLSSCTKTERLPSSDPLAELRKAPVNCSTAVCORIQCDIP 1020
Db 977 RLNOTVIMDRPQVTFSENLSSCTKTERLPSSDPLAELRKAPVNCSTAVCORIQCDIP 1036
Qy 1021 FFGIOBEFNATLKGNSLSPWYIKTNSHLLIVSTAEILFENDSVFTLLPQOGAFVRSQTET 1080
Db 1037 FFGIOBEFNATLKGNSLSPWYIKTNSHLLIVSTAEILFENDSVFTLLPQOGAFVRSQTET 1096
Qy 1081 KVEPEVNPPLPLVGVSSVGGHLLALITAAALYKLGFFKQYKDMWSEGGPPGASBPQ 1137
Db 1097 KVEPEVNPPLPLVGVSSVGGHLLALITAAALYKLGFFKQYKDMWSEGGPPGASBPQ 1153

RESULT 5

US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vlieten, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,899
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 278666/32684
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-605-672-3

Query Match
Best Local Similarity 99.5%; Score 5852; DB 2; Length 1153;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDENAMTPQENARFGQSVVQLQGSVVVVGAPQBIIVAAQORGSLYQCDYSTGSCPEI 60
Db 17 FNLDENAMTPQENARFGQSVVQLQGSVVVVGAPQBIIVAAQORGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVAVNMSLGLSLAATTPPOLLACGPTVHOTCSENTYVKGCLCFGLFOSNLROQPOK 120
Db 77 RLQVPVAVNMSLGLSLAATTPPOLLACGPTVHOTCSENTYVKGCLCFGLFOSNLROQPOK 136
Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMKEWSTVMBQKKSKTSLSLMOYSSEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMKEFVSTVMBQKKSKTSLSLMOYSSEF 196
Qy 181 RIHETFEFQNNPNSRIKPIITOLLGRTHATGLRKVVRELFNITNGARKNAFKILPFL 240
Db 197 RIHETFEFQNNPNSRLVKPIITOLLGRTHATGLRKVVRELFNITNGARKNAFKILVVI 256
Qy 241 TDGEKFGDPLGYEDVIPELDREGVIRVYVIGDAPFRSEKSRQELNTVASKPRDHVFQIN 300
Db 257 TDGEKFGDPLGYEDVIPEDADREGVIRVYVIGDAPFRSEKSRQELNTIASKPRDHVFQIN 316
Qy 301 NFEALKTIONQIREKIPIAIEGTOTGSSSSPEHESQBGFSAAITSNGLLSLTVGSDVWAG 360
Db 317 NFEALKTIONQIREKIPIAIEGTOTGSSSSPEHESQBGFSAAITSNGLLSLTVGSDVWAG 376
Qy 361 GVFLYTSKEKSTFINMTVDSDMDNDAYLGAAAIILNRNVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVDSDMDNDAYLGAAAIILNRNVQSLVGLGAPRYOHIGLVAMFR 436
Qy 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVILGAPHYHQTGQGVSVLCP 480
Db 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVILGAPHYHQTGQGVSVLCP 496
Qy 481 PRGQARWQCDAYLVGEQGPWGRFGAALTVDLVGVNGDKLTDVAIGAPGEEDNRGAVILF 540
Db 497 PRGQARWQCDAYLVGEQGPWGRFGAALTVDLVGVNGDKLTDVAIGAPGEEDNRGAVILF 556
Qy 541 HGTSGGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTGVAQHVLRLRSQ 600
Db 557 HGTSGGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTGVAQHVLRLRSQ 616
Qy 601 PVLRVKATMEFNPREVARNVFECDQVVKGEAGSVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKATMEFNPREVARNVFECDQVVKGEAGSVRVCLHVQKSTRDLREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFGNLRVLAEDAQRULTALFPEKKGNDNICODDLSITPFSMSLDCLVVG 780
Db 737 SLVGTPLSAFGNLRVLAEDAQRULTALFPEKKGNDNICODDLSITPFSMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDGDSYKQWTFPPDLDSYRKVSTIQNRSQSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGDSYKQWTFPPDLDSYRKVSTIQNRSQSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTENITPDVDSKASLGKLLKXAVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTENITPDVDSKASLGKLLKXAVTSENMPRTNKTEF 916
Qy 901 QLELPVKYAVVMVTVSHGVSTKYLANFTASENTSRVMQHOYQVSNLQORSPLISLVLFPV 960
Db 917 QLELPVKYAVVMVTVSHGVSTKYLANFTASENTSRVMQHOYQVSNLQORSPLISLVLFPV 976
Qy 961 RLNOTVIMDRPQVTFSENLSSCTKTERLPSSDPLAELRKAPVNCSTAVCORIQCDIP 1020
Db 977 RLNOTVIMDRPQVTFSENLSSCTKTERLPSSDPLAELRKAPVNCSTAVCORIQCDIP 1036

QY 1021 FFGIOEFNATLKGNSLSPWYIKTSHNLLIYSTABILENDVSFTLLPGQAFVRSOTET 1080
DB 1037 FFGIOEFNATLKGNSLSPWYIKTSHNLLIYSTABILENDVSFTLLPGQAFVRSOTET 1096
QY 1081 KVEPFVFNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1137
DB 1097 KVEPFVFNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1153

RESULT 6
US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: van der Vieren, Monica
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-482-293A-3

Query Match 99.5%; Score 5852; DB 2; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNDLTENAMTFQENARGFGQSVVQLOGSRVVVVGAPQRIYAANORGSLYQCDYSTGSCBPI 60
DB 17 FNDLTENAMTFQENARGFGQSVVQLOGSRVVVVGAPQRIYAANORGSLYQCDYSTGSCBPI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACQPTVHQTCSENTYVKGCFLFGSNLRQPOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACQPTVHQTCSENTYVKGCFLFGSNLRQPOK 136

QY 121 FPEALRGCPQEDSDIAFLIDSGSIIIPDPRMKEFVSTWMEQLKSKTFLSLMQYSEEP 180

DB 137 FPEALRGCPQEDSDIAFLIDSGSIIIPDPRMKEFVSTWMEQLKSKTFLSLMQYSEEP 196
QY 181 RIHFTFEFQNNPRLSIKPIITQLGRTHATGLRKVKVRELFNITNGARKVAKILPLL 240
DB 197 RIHFTFEFQNNPRLSIKPIITQLGRTHATGLRKVKVRELFNITNGARKVAKILVVI 256
QY 241 TDGKFGDPLGYEDVPELDRGVIRVYVVGDAFRSEKSRQELNTVASKPEPRDHVFQIN 300
DB 257 TDGKFGDPLGYEDVPELDRGVIRVYVVGDAFRSEKSRQELNTVASKPEPRDHVFQIN 316
QY 301 NFEALKTIQNLREKIPAEIGTOTGSSSSPHEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIQNLREKIPAEIGTOTGSSSSPHEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
QY 361 GVELYTSKSKSTFINMTRVDSMDNDAYLGAAAILNRRVQSLVGLGAPRYOHIGLVAMER 420
DB 377 GVELYTSKSKSTFINMTRVDSMDNDAYLGAAAILNRRVQSLVGLGAPRYOHIGLVAMER 436
QY 421 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVVCP 480
DB 437 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVVCP 496
QY 481 PRGQARWQCDVLYGEGQDPWGRFGAALTVDVNGDKLTDVAIGAPGEDNRCAYLF 540
DB 497 PRGQARWQCDVLYGEGQDPWGRFGAALTVDVNGDKLTDVAIGAPGEDNRCAYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSRLQYFGQSLGSGQDLTMDGLVDLTVGAQHVLRLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSRLQYFGQSLGSGQDLTMDGLVDLTVGAQHVLRLRSQ 616
QY 601 PVLRYKALMEFNPREVARNFECDQVYKGEAGEVRVCLHVOKSTRDLRREGQOSVVT 660
DB 617 PVLRYKALMEFNPREVARNFECDQVYKGEAGEVRVCLHVOKSTRDLRREGQOSVVT 676
QY 661 YDLALDSGRPHSRVFNETKSTTRQTOVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPHSRVFNETKSTTRQTOVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAFGNLRPLAEDAQRLFTALPFPEKKGNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFGNLRPLAEDAQRLFTALPFPEKKGNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSIVTNIITPDVDSKASLGNKLLKANVTSENNMPTNKTEP 900
DB 857 SGALKSTSCSINHPIFPENSIVTNIITPDVDSKASLGNKLLKANVTSENNMPTNKTEP 916
QY 901 QLELPVKYAVVMVYTSKGVSTKYLNFASNTSRVMQHQYQVSNLQSRSLPLSLVFLVPV 960
DB 917 QLELPVKYAVVMVYTSKGVSTKYLNFASNTSRVMQHQYQVSNLQSRSLPLSLVFLVPV 976
QY 961 RLNQTVIWDPRPQVTFPSENLSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP 1020
DB 977 RLNQTVIWDPRPQVTFPSENLSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP 1036
QY 1021 FPGIOEFNATLKGNSLSPWYIKTSHNLLIYSTABILENDVSFTLLPGQAFVRSOTET 1080
DB 1037 FPGIOEFNATLKGNSLSPWYIKTSHNLLIYSTABILENDVSFTLLPGQAFVRSOTET 1096
QY 1081 KVEPFVFNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1137
DB 1097 KVEPFVFNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1153

RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/943,363
FILING DATE:
CLASSIFICATION: 530
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

Query Match 99.5%; Score 5852; DB 2; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY	1	FNLDTENAMTFQENARFGQSVVQLOGSRVWVGAPQEIYAANQORGLYOCDSYTGSCBPI	60
DB	17	FNLDTENAMTFQENARFGQSVVQLOGSRVWVGAPQEIYAANQORGLYOCDSYTGSCBPI	76
QY	61	RLQVPVEAVNMVSLGLSLAATTGPPQLLACGPTVHOTCSNTYVKGCLFLFGNLRQOPQK	120
DB	77	RLQVPVEAVNMVSLGLSLAATTGPPQLLACGPTVHOTCSNTYVKGCLFLFGNLRQOPQK	136
QY	121	FPEALRGCEQSDIAFLIDGSGSIIIPDFRMEKQVSVTMEQLKSKTFLPSIMOYSEFP	180
DB	137	FPEALRGCEQSDIAFLIDGSGSIIIPDFRMEKQVSVTMEQLKSKTFLPSIMOYSEFP	196
QY	181	RLHFTKEQNNPNSLKPITOLGRTHATGLRKVVRELFTNNGARKNAFKILFLL	240
DB	197	RLHFTKEQNNPNSLKPITOLGRTHATGLRKVVRELFTNNGARKNAFKILVVI	256
QY	241	TGCEKFGDPLGYEDVPIPELDREGVIRYIVGVGDAPFSEKSRQELNVTASKPPRDVDFQIN	300
DB	257	TGCEKFGDPLGYEDVPIPELDREGVIRYIVGVGDAPFSEKSRQELNVTASKPPRDVDFQIN	316
QY	301	NFEALATQONREKIPALEGTQGTSSSFHEMSQEGFSAATNSGPLLSTVGSVDWAG	360
DB	317	NFEALATQONREKIPALEGTQGTSSSFHEMSQEGFSAATNSGPLLSTVGSVDWAG	376

QY	361	GVFLYTSKEKSTFIMTRVDSQNDAYLGAAAAIILNRVQSLVLTGAPRYQHIGLVAMFR	420
DB	377	GVFLYTSKEKSTFIMTRVDSQNDAYLGAAAAIILNRVQSLVLTGAPRYQHIGLVAMFR	436
QY	421	QNTGMWESNANVKCTQICAGYFAGSLCSVDVDSNGSTDLVLICAGPHYYPQTRGGQSVQVCP	480
DB	437	QNTGMWESNANVKCTQICAGYFAGSLCSVDVDSNGSTDLVLICAGPHYYPQTRGGQSVQVCP	496
QY	481	PRGORARWQCDVLYGBQGPWGRFGAALTVDVNGDKLTDVAIGAPGEDNRCGAVTLF	540
DB	497	PRGORARWQCDVLYGBQGPWGRFGAALTVDVNGDKLTDVAIGAPGEDNRCGAVTLF	556
QY	541	HGTSGSGISPSHSQRIAGSKLSPLRQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ	600
DB	557	HGTSGSGISPSHSQRIAGSKLSPLRQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ	616
QY	601	PVLAVKALMEFNPREVARNFECNDQVVKGEAGEVRVCLAVOKSTRDLREGQIQSVVT	660
DB	617	PVLAVKALMEFNPREVARNFECNDQVVKGEAGEVRVCLAVOKSTRDLREGQIQSVVT	676
QY	661	YDLALDSGRPHSRVAFVNETKSTTRRQTVLGLTQCTETKLQLPNCIEDPVPVILRLNF	720
DB	677	YDLALDSGRPHSRVAFVNETKSTTRRQTVLGLTQCTETKLQLPNCIEDPVPVILRLNF	736
QY	721	SLVGTPLSAPGNLRPVLAEQAQLFTALPPPEKNCNDNICODDLSITFSFMSLDCLVYG	780
DB	737	SLVGTPLSAPGNLRPVLAEQAQLFTALPPPEKNCNDNICODDLSITFSFMSLDCLVYG	796
QY	781	GPREFNVTVVRNDGDSYRQTVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV	840
DB	797	GPREFNVTVVRNDGDSYRQTVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV	856
QY	841	SGALKSTSCSINHPIFPENSESVTNITPFDVDSKASLGNKLLKANVTSENMMPTNKTEP	900
DB	857	SGALKSTSCSINHPIFPENSESVTNITPFDVDSKASLGNKLLKANVTSENMMPTNKTEP	916
QY	901	QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMOHQYOVSNLQORSLSPLSLVFLVPV	960
DB	917	QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMOHQYOVSNLQORSLSPLSLVFLVPV	976
QY	961	RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCIP	1020
DB	977	RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCIP	1036
QY	1021	FFGQEBEFNATLKGNSLSPDWIKTSHNHLIVSTAELFNDSPVTLAPQCGAFVRSOTET	1080
DB	1037	FFGQEBEFNATLKGNSLSPDWIKTSHNHLIVSTAELFNDSPVTLAPQCGAFVRSOTET	1096
QY	1081	KVEPPEVNPPLIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ	1137
DB	1097	KVEPPEVNPPLIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ	1153

RESULT 8
US-09-193-043-3
; Sequence 3, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6251395e1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193,043
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03

; NUMBER OF SEQ ID NOS: 114									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 3									
; LENGTH: 1153									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-193-043-3									
Query Match 99.5%; Score 5852; DB 3; Length 1153;									
Best Local Similarity 99.2%; Pred. No. 0;									
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;									
Qy	1	FNLDTENAMTFOENARFGQSVVQLGSGSVVVGAPQEIIVAAQSGSLYQCDYSTGSCPEI	60						
Db	17	FNLDTENAMTFOENARFGQSVVQLGSGSVVVGAPQEIIVAAQSGSLYQCDYSTGSCPEI	76						
Qy	61	RLOQVPEAVNMSLGLSLAATTSPPQLLAGCTVHQTCSNTYVKGELCFPLGSLNLRQOPQK	120						
Db	77	RLOQVPEAVNMSLGLSLAATTSPPQLLAGCTVHQTCSNTYVKGELCFPLGSLNLRQOPQK	136						
Qy	121	FPEALRGCPQSDSDIAPLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTLPFLMOYSEEF	180						
Db	137	FPEALRGCPQSDSDIAPLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTLPFLMOYSEEF	196						
Qy	181	RIHPTFKFQNNPNRSLIKPIIOLGRTHATGLRKVRELFTNITNGARKNAFKILFL	240						
Db	197	RIHPTFKFQNNPNRSLIKPIIOLGRTHATGLRKVRELFTNITNGARKNAFKILFL	256						
Qy	241	TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFRSEKSRQELNTVASKPPRDHVFQIN	300						
Db	257	TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFRSEKSRQELNTVASKPPRDHVFQIN	316						
Qy	361	GVFLYTSKEKSTINTRVDSMDNDAYLGAAAILLRNVQSLVLCAPRYQHIGLVAMFR	420						
Db	377	GVFLYTSKEKSTINTRVDSMDNDAYLGAAAILLRNVQSLVLCAPRYQHIGLVAMFR	436						
Qy	421	QNTGMESNANVGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYETRGQSVSCPL	480						
Db	437	QNTGMESNANVGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYETRGQSVSCPL	496						
Qy	481	PRQORARWQCDVLYGEQGPWRFGAALTVDVNGDKLTDVAICAPGEENRGAVILF	540						
Db	497	PRQORARWQCDVLYGEQGPWRFGAALTVDVNGDKLTDVAICAPGEENRGAVILF	556						
Qy	541	HGTSGSGIGPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLDVLTVGAQGHVILLRSQ	600						
Db	557	HGTSGSGIGPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLDVLTVGAQGHVILLRSQ	616						
Qy	601	PVLVRKALMEFNPREVARNVFCNDQVVKGEAGEVRVCLHYOKSTRDRLRGQIQSVVT	660						
Db	617	PVLVRKALMEFNPREVARNVFCNDQVVKGEAGEVRVCLHYOKSTRDRLRGQIQSVVT	676						
Qy	661	YDLALDSGRPHSAFVNETKNSRROTQVLGLTQTCETTLKQLPNCIEDPVSPIVLRNF	720						
Db	677	YDLALDSGRPHSAFVNETKNSRROTQVLGLTQTCETTLKQLPNCIEDPVSPIVLRNF	736						
Qy	721	SLVGTPLSAPGNTLPVLAEDAQRLFTALPPEKNCNDNTICQDDESITFSFMSLCLVVG	780						
Db	737	SLVGTPLSAPGNTLPVLAEDAQRLFTALPPEKNCNDNTICQDDESITFSFMSLCLVVG	796						
Qy	781	GPREFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV	840						
Db	797	GPREFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV	856						
Qy	841	SGALKSTSCSINPIPIPENSEVTFNTFDVDSKASLGKLLKANVTSENMPRTNKTEF	900						
Db	857	SGALKSTSCSINPIPIPENSEVTFNTFDVDSKASLGKLLKANVTSENMPRTNKTEF	916						
Qy	901	QLELPVKYAVYVVTSHGVSTKYNLTASNTSRVMQHOYQVSNLQORSLPISLVFLVPV	960						

Db	917	QLELPVKYAVYVVTSHGVSTKYNLTASNTSRVMQHOYQVSNLQORSLPISLVFLVPV	976						
Qy	961	RLNCTVIMDRPOVTFSENLSTCHTKERLPSHSDFLAELRKAPVVCNIAVCORIQCDDIP	1020						
Db	977	RLNCTVIMDRPOVTFSENLSTCHTKERLPSHSDFLAELRKAPVVCNIAVCORIQCDDIP	1036						
Qy	1021	PFQIEEFENATLKGNSLSDWYIKTSHNHLIVSTABILFNDVSFLLPQCGAFVRSQTEI	1080						
Db	1037	PFQIEEFENATLKGNSLSDWYIKTSHNHLIVSTABILFNDVSFLLPQCGAFVRSQTEI	1096						
Qy	1081	KVEPEFVNPLPLVGVSSVGLLILALITAAALYKLGPEKROYKMMSEGGPPGABPQ	1137						
Db	1097	KVEPEFVNPLPLVGVSSVGLLILALITAAALYKLGPEKROYKMMSEGGPPGABPQ	1153						
RESULT 9									
US-09-688-307A-3									
; Sequence 3, Application US/09688307A									
; Patent No. 6432404									
; GENERAL INFORMATION:									
; APPLICANT: Gallatin, Michael W.									
; APPLICANT: Van der Vieren, Monica									
; TITLE OF INVENTION: No. 6432404el Human Beta-2									
; FILE REFERENCE: 27866/36646									
; CURRENT APPLICATION NUMBER: US/09/688,307A									
; CURRENT FILING DATE: 2000-10-13									
; PRIOR APPLICATION NUMBER: 09/193,043									
; PRIOR FILING DATE: 1998-11-16									
; PRIOR APPLICATION NUMBER: 08/605,672									
; PRIOR FILING DATE: 1996-02-22									
; PRIOR APPLICATION NUMBER: 08/173,497									
; PRIOR FILING DATE: 1993-12-23									
; PRIOR APPLICATION NUMBER: 08/286,889									
; PRIOR FILING DATE: 1994-08-05									
; PRIOR APPLICATION NUMBER: 08/362,652									
; PRIOR FILING DATE: 1994-12-21									
; PRIOR APPLICATION NUMBER: 08/943,363									
; PRIOR FILING DATE: 1997-10-03									
; NUMBER OF SEQ ID NOS: 114									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 3									
; LENGTH: 1153									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-688-307A-3									
Query Match 99.5%; Score 5852; DB 4; Length 1153;									
Best Local Similarity 99.2%; Pred. No. 0;									
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;									
Qy	1	FNLDTENAMTFOENARFGQSVVQLGSGSVVVGAPQEIIVAAQSGSLYQCDYSTGSCPEI	60						
Db	17	FNLDTENAMTFOENARFGQSVVQLGSGSVVVGAPQEIIVAAQSGSLYQCDYSTGSCPEI	76						
Qy	61	RLOQVPEAVNMSLGLSLAATTSPPQLLAGCTVHQTCSNTYVKGELCFPLGSLNLRQOPQK	120						
Db	77	RLOQVPEAVNMSLGLSLAATTSPPQLLAGCTVHQTCSNTYVKGELCFPLGSLNLRQOPQK	136						
Qy	121	FPEALRGCPQSDSDIAPLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTLPFLMOYSEEF	180						
Db	137	FPEALRGCPQSDSDIAPLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTLPFLMOYSEEF	196						
Qy	181	RIHPTFKFQNNPNRSLIKPIIOLGRTHATGLRKVRELFTNITNGARKNAFKILFL	240						
Db	197	RIHPTFKFQNNPNRSLIKPIIOLGRTHATGLRKVRELFTNITNGARKNAFKILFL	256						
Qy	241	TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFRSEKSRQELNTVASKPPRDHVFQIN	300						
Db	257	TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFRSEKSRQELNTVASKPPRDHVFQIN	316						
Qy	301	NFEALKTIONLREKIPAEIGTQSGSSSFHEMSQGFSAITNSGFLPSTVGSYDWAG	360						

317 NFEALKTIONQREKI PAIEGTQTGSSSSPEHEMSQEGFSAATNSNGPILLSSTVGSYDWAG 376
361 GVFLYTSKEKSTFNMTVRVDSMDNDAYLGAAAIILNRVQSIVLGAAPRYOHIGLVAMPR 420
377 GVFLYTSKEKSTFNMTVRVDSMDNDAYLGAAAIILNRVQSIVLGAAPRYOHIGLVAMPR 436
421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQSVNCP 480
437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQSVNCP 496
481 PRGORARWQCDVAVLYGEGQPGWRFGAALTVDVNGDKLTDVAIGAPGEDNNGAVYLF 540
497 PRGORARWQCDVAVLYGEGQPGWRFGAALTVDVNGDKLTDVAIGAPGEDNNGAVYLF 556
541 HGTSGSGISPHSQRISAGSKLSPRLQYEGQISGGQDLTMGLVDLTVGAQGHVLLRSQ 600
557 HGTSGSGISPHSQRISAGSKLSPRLQYEGQISGGQDLTMGLVDLTVGAQGHVLLRSQ 616
601 PVLRVKALMEFNPREVARNVFECDQVVKGEAGVRVCLHVOKSTRDLRREGQIQSVVT 660
617 PVLRVKALMEFNPREVARNVFECDQVVKGEAGVRVCLHVOKSTRDLRREGQIQSVVT 676
661 YDLALDSGRPHSRVFNKSTRQTOVLGTTCTETKLQLPNCIEDPVSPIVLRNF 720
677 YDLALDSGRPHSRVFNKSTRQTOVLGTTCTETKLQLPNCIEDPVSPIVLRNF 736
721 SLVGTPLSAFAGNLRPVLAEDAQRFLTALPFPEKNCNDNI CODDLSITFSFMSDCLVVG 780
737 SLVGTPLSAFAGNLRPVLAEDAQRFLTALPFPEKNCNDNI CODDLSITFSFMSDCLVVG 796
781 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYKRVSTLQNRORSWRACESASSTEV 840
797 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYKRVSTLQNRORSWRACESASSTEV 856
841 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900
857 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 916
901 QLELPVKAVMVTSHGVSTKYLNFASNTSRVQOYQVSNLQORSWRACESASSTEV 960
917 QLELPVKAVMVTSHGVSTKYLNFASNTSRVQOYQVSNLQORSWRACESASSTEV 976
961 RLNQTVIWDROPVTPSENLSSTCHTKERLPSSHDFLAELKAPVNCISIAVCQRIQCDIP 1020
977 RLNQTVIWDROPVTPSENLSSTCHTKERLPSSHDFLAELKAPVNCISIAVCQRIQCDIP 1036
1021 PFGIQEEFNATLKGNSLSPDWIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQDET 1080
1037 PFGIQEEFNATLKGNSLSPDWIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQDET 1096
1081 KVEPEVENPLPLIVGSSVGGILLALITAALYKLGPFKROYKDMWSBGGPPGAEPO 1137
1097 KVEPEVENPLPLIVGSSVGGILLALITAALYKLGPFKROYKDMWSBGGPPGAEPO 1153

RESULT 10
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. 6620915
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6620915el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652

; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3
Query Match 99.5%; Score 5852; DB 4; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVVCAPQBIIVAAQNRGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVVCAPQBIIVAAQNRGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATSPOLLACQPTVHQTCSNTYVKGCLFLFGSNLRQOPQOK 120
DB 77 RLQVPVEAVNMSLGLSLAATSPOLLACQPTVHQTCSNTYVKGCLFLFGSNLRQOPQOK 136
QY 121 FPALRCPCPEDSDIAPLIDSGSIIIPHDPRRMKEMVSTVMBOLKSKTLPFLMLOYSEEP 180
DB 137 FPALRCPCPEDSDIAPLIDSGSIIIPHDPRRMKEMVSTVMBOLKSKTLPFLMLOYSEEP 196
QY 181 RIHFTPEKQNNPNRSLIKPIITQLGRHTATGLRKRVRELFNITNGARKNAFKILVVI 240
DB 197 RIHFTPEKQNNPNRSLIKPIITQLGRHTATGLRKRVRELFNITNGARKNAFKILVVI 256
QY 241 TDCEKCPDGYEDVTPELDREGVIRVYGVGDAFRSEKSRQELNTVASKPRDHVQIN 300
DB 257 TDCEKCPDGYEDVTPELDREGVIRVYGVGDAFRSEKSRQELNTVASKPRDHVQIN 316
QY 301 NFZALKTIQNLREKI PAIEGTQTGSSSSPEHEMSQEGFSAATNSNGPILLSSTVGSYDWAG 360
DB 317 NFZALKTIQNLREKI PAIEGTQTGSSSSPEHEMSQEGFSAATNSNGPILLSSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFNMTVRVDSMDNDAYLGAAAIILNRVQSIVLGAAPRYOHIGLVAMPR 420
DB 377 GVFLYTSKEKSTFNMTVRVDSMDNDAYLGAAAIILNRVQSIVLGAAPRYOHIGLVAMPR 436
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQSVNCP 480
DB 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQSVNCP 496
QY 481 PRGORARWQCDVAVLYGEGQPGWRFGAALTVDVNGDKLTDVAIGAPGEDNNGAVYLF 540
DB 497 PRGORARWQCDVAVLYGEGQPGWRFGAALTVDVNGDKLTDVAIGAPGEDNNGAVYLF 556
QY 541 HGTSGSGISPHSQRISAGSKLSPRLQYEGQISGGQDLTMGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPHSQRISAGSKLSPRLQYEGQISGGQDLTMGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLRVKALMEFNPREVARNVFECDQVVKGEAGVRVCLHVOKSTRDLRREGQIQSVVT 660
DB 617 PVLRVKALMEFNPREVARNVFECDQVVKGEAGVRVCLHVOKSTRDLRREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVFNKSTRQTOVLGTTCTETKLQLPNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPHSRVFNKSTRQTOVLGTTCTETKLQLPNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAFAGNLRPVLAEDAQRFLTALPFPEKNCNDNI CODDLSITFSFMSDCLVVG 780
DB 737 SLVGTPLSAFAGNLRPVLAEDAQRFLTALPFPEKNCNDNI CODDLSITFSFMSDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYKRVSTLQNRORSWRACESASSTEV 840
DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYKRVSTLQNRORSWRACESASSTEV 856
QY 841 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900

Db 857 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTET 916
Qy 901 QLELPVKYAVYVMTSHGVSSTYLAFTASENTRVMOHQYOVSNLQORSLSPLSLVPLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSSTYLAFTASENTRVMOHQYOVSNLQORSLSPLSLVPLVPV 976
Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCQRIQCDDIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCQRIQCDDIP 1036
Qy 1021 PFGIOEEFNATLKNLSFDWIKTSHNHLIYSTABILFNDSVFTLLPQGGAFVRSQDET 1080
Db 1037 PFGIOEEFNATLKNLSFDWIKTSHNHLIYSTABILFNDSVFTLLPQGGAFVRSQDET 1096
Qy 1081 KVEPPEVPNPLPLIVGSSVGGLLLALITAALYKLGFFKROQKDMSEGGPPGABPQ 1137
Db 1097 KVEPPEVPNPLPLIVGSSVGGLLLALITAALYKLGFFKROQKDMSEGGPPGABPQ 1153

RESULT 11

US-08-476-062A-43
; Sequence 43, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Annaut, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212,573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal

US-08-476-062A-43

Query Match 98.94; Score 5821.5; DB 2; Length 1152;
Best Local Similarity 98.94; Pred. No. 0;
Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFOENARGFGOSVVOLQSGRVVVGAPQBIIVAAVNOQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGOSVVOLQSGRVVVGAPQBIIVAAVNOQSGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVAVVNMVSLGSLAAATTPPOLLAGCPTVHQTCSNTYVKGCLFELPQSNLRQOPQK 120
Db 77 RLQVPVAVVNMVSLGSLAAATTPPOLLAGCPTVHQTCSNTYVKGCLFELPQSNLRQOPQK 136
Qy 121 PPEALRGCPQSDSDIAFLIDGSGSIIIPDFFRMKEWSTVMEQLKSKSKTLPFLMQYSSEF 180
Db 137 PPEALRGCPQSDSDIAFLIDGSGSIIIPDFFRMKEWSTVMEQLKSKSKTLPFLMQYSSEF 196
Qy 181 RIHTTFKFPQNNPNSRLIIPITOLLGRTHATGLRKVVRBELFNI TNGARKNAKILPLL 240
Db 197 RIHTTFKFPQNNPNSRLIIPITOLLGRTHATGLRKVVRBELFNI TNGARKNAKILPLL 256
Qy 241 TDGKFGDPLGYEDVPELDRGVRVYVIGVDAPFRSEKSRQELNLTVAASKPRDRHVFQIN 300
Db 257 TDGKFGDPLGYEDVPELDRGVRVYVIGVDAPFRSEKSRQELNLTVAASKPRDRHVFQIN 316
Qy 301 NFEALKTIQNLQRKIPAEIGCTGTGSSSSPHEMSQSGFSAALITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIQNLQRKIPAEIGCTGTGSSSSPHEMSQSGFSAALITNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
Qy 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEYQTRGGQVSVCP 480
Db 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEYQTRGGQVSVCP 496
Qy 481 PRGORARWQCDVLYGEOGPWGHFGAALTVLGDVNGDKLTDVAIGAPGEDNDRGAVLP 540
Db 497 PRG-RARWQCDVLYGEOGPWGHFGAALTVLGDVNGDKLTDVAIGAPGEDNDRGAVLP 555
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGQDLTMDGLVDLTVGAQCHVLLLRSQ 600
Db 556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGQDLTMDGLVDLTVGAQCHVLLLRSQ 615
Qy 601 PVLRVKAIEMFNPREVARNVFECDVQVYKGEAGEVRLVOKSTRDLREGQIQSVVT 660
Db 616 PVLRVKAIEMFNPREVARNVFECDVQVYKGEAGEVRLVOKSTRDLREGQIQSVVT 675
Qy 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 676 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 735
Qy 721 SLVGTPLSAPGNLAPVLAEDAQRFTALPPEKKGNDNICODDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAPGNLAPVLAEDAQRFTALPPEKKGNDNICODDLSITFSFMSLDCLVVG 795
Qy 781 GPREPNVTVVRNDGEDSYRTQVTFPPPLDLSYKRVSTLQNRQSRWSRLACESASSTEV 840
Db 796 GPREPNVTVVRNDGEDSYRTQVTFPPPLDLSYKRVSTLQNRQSRWSRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTET 900
Db 856 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTET 915
Qy 901 QLELPVKYAVYVMTSHGVSSTYLAFTASENTRVMOHQYOVSNLQORSLSPLSLVPLVPV 960
Db 916 QLELPVKYAVYVMTSHGVSSTYLAFTASENTRVMOHQYOVSNLQORSLSPLSLVPLVPV 975
Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCQRIQCDDIP 1020
Db 976 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCQRIQCDDIP 1035
Qy 1021 PFGIOEEFNATLKNLSFDWIKTSHNHLIYSTABILFNDSVFTLLPQGGAFVRSQDET 1080
Db 1036 PFGIOEEFNATLKNLSFDWIKTSHNHLIYSTABILFNDSVFTLLPQGGAFVRSQDET 1095
Qy 1081 KVEPPEVPNPLPLIVGSSVGGLLLALITAALYKLGFFKROQKDMSEGGPPGABPQ 1137

Db 1096 KVEPFEVNPPLIVGSSVGGLLALLALITAAALYKLGFFRQYKDMSEGPPGAEPQ 1152
|||||
RESULT 12
PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaut
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8306
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US96-01314-43
Query Match 98.9%; Score 5821.5; DB 5; Length 1152;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;
QY 1 FNLDTENAMTFOENARGCQSVVVOGGSRVAVGAPQEIIVAAVNOGSLYCCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGCQSVVVOGGSRVAVGAPQEIIVAAVNOGSLYCCDYSTGSCPEI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFLGSLNLRQPOK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFLGSLNLRQPOK 136
QY 121 PEARLGCQPSDIAFLIDGSGSIIPDRFMKWSVSTWEOIKSKTLPSLMQYSEEP 180
DB 137 PEARLGCQPSDIAFLIDGSGSIIPDRFMKWSVSTWEOIKSKTLPSLMQYSEEP 196
QY 181 RHFTPKFQNNPNRSLKPTQLLGRTHATGLRKVVRELFNITNGARKNAKFLIPL 240
DB 197 RHFTPKFQNNPNRSLKPTQLLGRTHATGLRKVVRELFNITNGARKNAKFLIPL 256
QY 241 TDGKPGDPLGYEDVIPLEDRGVIRYVIGVDGAFRSEKSRQELMTVASKPPRDHVFQIN 300
DB 257 TDGKPGDPLGYEDVIPLEDRGVIRYVIGVDGAFRSEKSRQELMTVASKPPRDHVFQIN 316
QY 301 NFEALKTIONOLREKIFALEGTQTCSSSFEHMSQEGSAAITSNGLPSTVGSYDNAG 360
DB 317 NFEALKTIONOLREKIFALEGTQTCSSSFEHMSQEGSAAITSNGLPSTVGSYDNAG 376

QY 361 GVFLYTSKSKSTFINMTRVDSMDNDAYLGVAALILLNRVQSLVLAGAPRYOHIGLVAMER 420
DB 377 GVFLYTSKSKSTFINMTRVDSMDNDAYLGVAALILLNRVQSLVLAGAPRYOHIGLVAMER 436
QY 421 QNTGMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYTQRTGQGVSCPL 480
DB 437 QNTGMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYTQRTGQGVSCPL 496
QY 481 PRGORARWOCDAVLKXGQOPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNKGAVYLF 540
DB 497 PRG-RASWQCDAVLYGQGGQWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNKGAVYLF 555
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600
DB 556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVLDLTVGAQGHVLLRSQ 615
QY 601 PVLRVKALMEFNPREVARNVFECDQVYVKGKAGEVRVCLHVOKSTDRIRREGQIQSVT 660
DB 616 PVLRVKALMEFNPREVARNVFECDQVYVKGKAGEVRVCLHVOKSTDRIRREGQIQSVT 675
QY 661 YDLALDSGRPHSRVAFNETKNSRROTQVIGLQTCTETLKLQFPNCIEDPVSPTVLRINF 720
DB 676 YDLALDSGRPHSRVAFNETKNSRROTQVIGLQTCTETLKLQFPNCIEDPVSPTVLRINF 735
QY 721 SLVGTPLSAFQNLRLPVLAEQAORLFTALPPFKKNCNDNICODDLSTTFPSMSLDCLVVG 780
DB 736 SLVGTPLSAFQNLRLPVLAEQAORLFTALPPFKKNCNDNICODDLSTTFPSMSLDCLVVG 795
QY 781 GPREFNVTIVRNDGDSYRTQVTFPPPLDLYRKVSTLQNSORSWRLLACESASSTEV 840
DB 796 GPREFNVTIVRNDGDSYRTQVTFPPPLDLYRKVSTLQNSORSWRLLACESASSTEV 855
QY 841 SGALKSTCSINHPPIPPENSEVFNITFDVDSKASLQNLKLLKANVTSENNMPRTNTEF 900
DB 856 SGALKSTCSINHPPIPPENSEVFNITFDVDSKASLQNLKLLKANVTSENNMPRTNTEF 915
QY 901 QLELPVKYAVYVVTSHGVSTKYLNFASNTSRVMOHQYQVSNLQORSPLISVFLVPV 960
DB 916 QLELPVKYAVYVVTSHGVSTKYLNFASNTSRVMOHQYQVSNLQORSPLISVFLVPV 975
QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
DB 976 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1035
QY 1021 PFGIOEFNATLKNLSFDNVIKTSNHELLIVSTABILLFNDVSPTLLPGOGAFVRSOTET 1080
DB 1036 PFGIOEFNATLKNLSFDNVIKTSNHELLIVSTABILLFNDVSPTLLPGOGAFVRSOTET 1095
QY 1081 KVSEPEVNPPLIVGSSVGGLLALLALITAAALYKLGFFRQYKDMSEGPPGAEPQ 1137
DB 1096 KVSEPEVNPPLIVGSSVGGLLALLALITAAALYKLGFFRQYKDMSEGPPGAEPQ 1152

RESULT 13
5424399-2
; Patent No. 5424399
; APPLICANT: ARNAUT, M. AMIN
; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/78,871
; FILING DATE: 16-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 212,573
; FILING DATE: 28-JUN-1988
; SEQ ID NO: 2:
; LENGTH: 1152
5424399-2

Query Match 98.9%; Score 5821.5; DB 6; Length 1152;

Best Local Similarity 98.9%; Pred. No. 0;									
Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;									
Qy	1	FNLDTENAMTFOENARGFGQSVVOLGSRVVGAPQEIIVAAANQSGSLYOCYSTGSCBPI	60						
Db	17	FNLDTENAMTFOENARGFGQSVVOLGSRVVGAPQEIIVAAANQSGSLYOCYSTGSCBPI	76						
Qy	61	RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCPLFGSNLRQPOK	120						
Db	77	RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCPLFGSNLRQPOK	136						
Qy	121	FPFALGCCPOEDSDIAFLIDGSGSIIPHDPRMKWVSTVMEOLKSKTLFSLMOYSEEP	180						
Db	137	FPFALGCCPOEDSDIAFLIDGSGSIIPHDPRMKWVSTVMEOLKSKTLFSLMOYSEEP	196						
Qy	181	RIHFTFEKFNPNPRSLKIPITQLGRTHATGLAKVVRLEFNIITNGARKNAFKILFL	240						
Db	197	RIHFTFEKFNPNPRSLKIPITQLGRTHATGLAKVVRLEFNIITNGARKNAFKILFL	256						
Qy	241	TDGEKGDPLGYEDVPELDBEGVIRYVIGVGDAFRSEKSRQELANTVASKPRDHVQIN	300						
Db	257	TDGEKGDPLGYEDVPELDBEGVIRYVIGVGDAFRSEKSRQELANTVASKPRDHVQIN	316						
Qy	301	NFEALKTIONLREKIFAIEGTGTGSSSPFHEMSQEGFSAITNSGPLLSTVGSYDMAG	360						
Db	317	NFEALKTIONLREKIFAIEGTGTGSSSPFHEMSQEGFSAITNSGPLLSTVGSYDMAG	376						
Qy	361	GVFLYTSKEKSTINFRVDSNDAYLVAAAIILNRVQSLVLTGAPYOHIGLVAMFR	420						
Db	377	GVFLYTSKEKSTINFRVDSNDAYLVAAAIILNRVQSLVLTGAPYOHIGLVAMFR	436						
Qy	421	QNTGMESNANVKGTOIGVFGASLCSVDVDSNGSTDVLIGAPHYVETRGQGVSCPL	480						
Db	437	QNTGMESNANVKGTOIGVFGASLCSVDVDSNGSTDVLIGAPHYVETRGQGVSCPL	496						
Qy	481	PRGORARWQCDAYLGEQGFWRPGALTVLGDVNGDKLTDVAIGAPQEDNRGAVYLF	540						
Db	497	PRG-RARWQCDAYLGEQGFWRPGALTVLGDVNGDKLTDVAIGAPQEDNRGAVYLF	555						
Qy	541	HGTSGSGISPSHSORLAGSKLSPLOVFGQSLGGQDLTMDGLVDTVGAQGHVLLRQ	600						
Db	556	HGTSGSGISPSHSORLAGSKLSPLOVFGQSLGGQDLTMDGLVDTVGAQGHVLLRQ	615						
Qy	601	PVLVRKAIMEFNPREVARNVFCNDQVVKQKEAGEVRVCLHVOKSTRDLRREGQIOSVT	660						
Db	616	PVLVRKAIMEFNPREVARNVFCNDQVVKQKEAGEVRVCLHVOKSTRDLRREGQIOSVT	675						
Qy	661	YDLALDSGRPHSRAVNETKSTRQTVLGLTOTCETLKLQLPNCIEDPVSIVLRLNF	720						
Db	676	YDLALDSGRPHSRAVNETKSTRQTVLGLTOTCETLKLQLPNCIEDPVSIVLRLNF	735						
Qy	721	SLVGTPLSAFNGNLRPVLAEDAQRLLFTALFPFEKNCNDNI CQDDLSTITFSFMSLDCLVG	780						
Db	736	SLVGTPLSAFNGNLRPVLAEDAQRLLFTALFPFEKNCNDNI CQDDLSTITFSFMSLDCLVG	795						
Qy	781	GPREFNVTVVRNDGDSYQTVTPPPPLDSLVRKUSTLQNRQSRWRLACASSTEV	840						
Db	796	GPRESNVTVVRNDGDSYQTVTPPPPLDSLVRKUSTLQNRQSRWRLACASSTEV	855						
Qy	841	SGALKSTSCSINRPIPPENSEVTFNITFDVDSKASLGNKLLKXANTSENMPRTNKTEF	900						
Db	856	SGALKSTSCSINRPIPPENSEVTFNITFDVDSKASLGNKLLKXANTSENMPRTNKTEF	915						
Qy	901	QLELPVKYAVVMVTSHGVTSKYLNFTASNTSRVMQHQVQVNSLQORSPLISLFLVPV	960						
Db	916	QLELPVKYAVVMVTSHGVTSKYLNFTASNTSRVMQHQVQVNSLQORSPLISLFLVPV	975						
Qy	961	RLMOTVLDWPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCORIQCDIP	1020						
Db	976	RLMOTVLDWPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCORIQCDIP	1035						
Qy	1021	FFGIQBEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQET	1080						

Query Match
Best Local Similarity 60.9%; Pred. No. 9.8e-290; Indels 6; Gaps 4;
Matches 687; Conservative 142; Mismatches 294

Qy	1	FNLDTENAMTFOENARGFGQSVVOLGSRVVGAPQEIIVAAANQSGSLYOCYSTGSCBPI	60
Db	20	FNLDTENAMTFOENARGFGQSVVOLGSRVVGAPQEIIVAAANQSGSLYOCYSTGSCBPI	79
Qy	61	RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCPLFGSNLRQPOK	120
Db	80	GLQVPPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCPLFGSNLRQPOK	137
Qy	121	FPFALGCCPOEDSDIAFLIDGSGSIIPHDPRMKWVSTVMEOLKSKTLFSLMOYSEEF	180
Db	138	LPVSRQECPRQODIVFLIDGSGSIIPHDPRMKWVSTVMEOLKSKTLFSLMOYSEEF	197
Qy	181	RIHFTFEKFNPNPRSLKIPITQLGRTHATGLAKVVRLEFNIITNGARKNAFKILFL	240

Qy	541	HGTSGGSI	SPSHS	ORLAGS	KLS	PRLOY	FGQSL	SGGDL	TMD	GLVD	LT	VGA	QGH	VLL	RSQ	600
Db	557	HGVL	GPSIS	SPSHS	ORLAGS	QLS	SRLOY	FGQAL	SGGDL	TQD	GLVD	LT	VGA	RGQ	VLL	TR
Qy	601	PVL	RVK	AIM	EFP	PRE	VARN	FE	CND	QV	VK	KE	AGE	VR	V	CH
Db	617	PVL	WVG	SM	QFIP	AEI	PR	SA	FE	CR	EQ	V	W	SE	Q	TL
Qy	661	YD	LAL	D	SG	R	PHS	RA	V	NE	T	K	N	STR	Q	T
Db	677	LD	LAL	D	GR	LS	PR	AT	FO	E	T	K	N	SR	LS	R
Qy	721	SL	VGT	P	L	S	A	F	G	N	R	P	V	L	A	E
Db	737	TL	V	G	K	P	L	L	A	F	N	R	P	M	L	A
Qy	781	GP	RE	F	N	V	T	V	R	D	G	E	S	Y	T	O
Db	797	SN	LE	NA	E	V	M	V	M	D	G	E	S	Y	T	I
Qy	841	SG	AL	K	T	S	C	S	I	N	H	P	I	P	P	E
Db	855	SQ	GT	W	T	S	C	R	I	N	H	L	I	F	R	G
Qy	901	Q	L	E	L	P	V	K	I	A	T	M	O	T	S	H
Db	915	Q	L	E	L	P	V	K	I	A	T	M	O	T	S	H
Qy	960	V	R	L	N	O	T	V	I	M	D	R	P	O	V	T
Db	975	V	E	L	N	O	E	A	V	M	D	V	E	S	H	
Qy	1020	P	F	G	I	O	E	E	F	N	A	T	L	K	N	
Db	1035	P	S	F	S	V	Q	E	L	D	F	T	L	K		
Qy	1080	T	K	V	E	P	R	E	V	P	N	P	L	I		
Db	1095	T	V	L	E	K	I	V	H	A	P	T	P			

Search completed: June 7, 2004, 17:19:23
 Job time : 22.3484 secs

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:05:16 ; Search time 14.559 Seconds
(without alignments)
7512.163 Million cell updates/sec

Title: US-09-902-481B-5
Perfect score: 5876
Sequence: 1 FNLDTENAMTQENARGFGQ.....PFRQYKMMSEGGPPGAEPPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78: *
1: Piri: *
2: Piri2: *
3: Piri3: *
4: Piri4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5862	99.8	1153	1 RWU51B	cell surface glyco
2	4470	76.1	1153	2 S00551	leukocyte surface
3	3479	59.2	1163	1 RWU1C	cell surface glyco
4	1553.5	26.4	1170	2 S03308	cell surface glyco
5	1538.5	26.2	1163	2 S03126	lymphocyte fuction
6	1153	19.6	1179	2 A53213	integrin alpha-8 c
7	1103.5	18.8	1151	2 A45226	integrin alpha-1 c
8	1095	18.5	1170	2 I45914	integrin alpha-2 s
9	1071	18.2	1178	2 S44142	VLA-2 protein homo
10	1068	18.2	1181	2 A33998	integrin alpha-2 c
11	1062	18.1	1180	2 A35854	integrin alpha-1 c
12	663	11.3	1039	2 A41131	lymphocyte-Peyer's
13	635	10.8	1038	2 S06046	integrin alpha-4 c
14	630	10.7	1035	2 I58409	integrin alpha-9 c
15	614.5	10.5	1041	2 T31437	integrin alpha-cha
16	579.5	9.9	1054	2 JC7294	alphan integrin -
17	572.5	9.7	1051	2 A35761	cell surface glyco
18	567.5	9.7	1053	2 I55534	VLA-3 alpha subuni
19	555.5	9.5	1053	2 S44250	integrin alpha-5 c
20	542.5	9.2	1034	2 A36108	integrin alpha-v c
21	535	9.1	1044	2 T10050	integrin alpha-v c
22	532	9.1	1049	2 A27079	fibronectin recept
23	532	9.1	1073	2 B36429	integrin alpha-6 c
24	531.5	9.0	1072	2 A38457	integrin alpha-6 c
25	530.5	9.0	1051	2 A40021	integrin alpha-3 c
26	526	9.0	1048	2 A27421	integrin alpha-6 c
27	525.5	8.9	1091	2 A41543	integrin alpha-6 c
28	517	8.8	1044	2 S16516	integrin alpha-8 c
29	496.5	8.4	1146	2 S40311	integrin - fruit f

30	496	8.4	1394	2 A29637	position-specific
31	495.5	8.4	1039	2 A34269	integrin alpha-2b
32	494	8.4	1137	2 JCS950	integrin alpha-7 c
33	490.5	8.3	1037	2 A60163	glycoprotein IIB -
34	488	8.3	1135	2 I61186	alpha-7 integrin -
35	486	8.3	126	2 B30892	leukocyte adhesion
36	480	8.0	1106	2 S38783	integrin alpha cha
37	467	7.9	1226	2 S44824	PS42.1 protein -
38	453	7.7	1045	2 S60571	integrin alpha v c
39	445.5	7.6	1139	2 S28277	hypothetical prote
40	423	7.2	1115	2 T09403	integrin alpha cha
41	418.5	7.1	1115	2 T09433	integrin alpha cha
42	391	6.7	764	2 I36916	glycoprotein IIB -
43	309.5	5.3	1086	2 I18523	integrin alpha cha
44	299	5.1	604	2 I36917	glycoprotein IIB -
45	298.5	5.1	272	2 A55348	integrin alpha-1 -

ALIGNMENTS

RESULT 1

RWU1B
cell surface glycoprotein CD11b precursor [validated] - human
N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein M
eukocyte integrin alpha chain; neutrophil adherence receptor alphas chain
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text change 31-Dec-2000
C:Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I58567
R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.
J. Biol. Chem. 263, 12403-12411, 1988
A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, C
B.
A:Reference number: A31108; MUID:88315033; PMID:2457584
A:Accession: A31108
A:Molecule type: mRNA
A:Residues: 1-1153 <OR>
A:Cross-references: GB:J03925; NID:G187284; PIDN:AAA59544.1; PID:G307148
A:Note: part of this sequence was confirmed by protein sequencing
R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.
J. Cell Biol. 106, 2153-2158, 1988
A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor 1
A:Reference number: A28915; MUID:88257215; PMID:2454931
A:Accession: A28915
A:Molecule type: mRNA
A:Residues: 1-499,501-965, 'P', 967-1153 <ARN>
A:Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:G186935; PIDN:AAA59
A:Note: the authors translated the codon TAC for residue 1129 as Thr
A:Note: part of this sequence, including the amino end of the mature protein, was confi
R:Shelley, C.S.; Arnaout, M.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991
A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally reg
A:Reference number: A41600; MUID:92073318; PMID:1683702
A:Accession: A41600
A:Molecule type: DNA
A:Residues: 1-9 <SHE>
A:Cross-references: GB:W76724; NID:G180018; PIDN:AAA58410.1; PID:G553215
R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988
A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesi
A:Reference number: A94193; MUID:88190151; PMID:2833753
A:Accession: A30892
A:Molecule type: mRNA
A:Residues: 917-1042 <AR2>
A:Cross-references: GB:M18044
R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989
A:Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence receptor
A:Reference number: A32218; MUID:89098893; PMID:2563162
A:Accession: A32218
A:Molecule type: mRNA
A:Residues: 9-1153 <HIC>
A:Cross-references: GB:J04145; NID:G189068; PIDN:AAA59903.1; PID:G386975

A;Note: part of this sequence was confirmed by protein sequencing
R;Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.P.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
n during evolution.
A;Reference number: A46526; PMID:93123748; PMID:8419480
A;Accession: A46526
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-499,501-1153 <FLB>
A;Cross-references: GB:S52227; NID:G263047; PIDN:AA24821.1; PID:G263049
A;Note: the last three bases of intron 13, CAG, are included in some but not all mature
R;Pierce, M.W.; Remold-O'Donnell, B.; Todd III, R.P.; Arnaout, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986
A;Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
A;Reference number: A90664; MUID:87076671; PMID:3539202
A;Accession: A26091
A;Molecule type: protein
A;Residues: 17-31 <PIE>
A;Experimental source: granulocytes
R;Pahl, H.L.; Rosemarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992
A;Title: Characterization of the myeloid-specific CD11b promoter.
A;Reference number: 152567; MUID:92144986; PMID:1346576
A;Accession: 152567
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-9 <RES>
A;Cross-references: GB:M84477; NID:G180184; PIDN:AA51960.1; PID:G553219
C;Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
C;Genetics:
A;Gene: GDB:ITGAM; CR3A
A;Cross-references: GDB:120599; OMIM:120980
A;Map position: 16p11.2-16p11.2
A;Note: promoter contains a GATA motif and two Sp1 consensus binding sites
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C;Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-1153/Product: cell surface glycoprotein CD11b #status experimental <WAT>
F;17-1153/Domain: extracellular #status predicted <EXT>
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA>
F;465-473/Region: calcium/magnesium binding #status predicted
F;530-538/Region: calcium/magnesium binding #status predicted
F;593-601/Region: calcium/magnesium binding #status predicted
F;1109-1134/Region: calcium/magnesium binding #status predicted <TM>
F;1135-1153/Domain: intracellular #status predicted <INT>
F;86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY	1	FNLDTENAMTFQENARGFQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI	60
DB	17	FNLDTENAMTFQENARGFQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI	76
QY	61	RLOPVVEAVNNSGLSLAATSPOLLACGTPVHOTCSNTYVKGCLFLFGSLNRQPOK	120
DB	77	RLOPVVEAVNNSGLSLAATSPOLLACGTPVHOTCSNTYVKGCLFLFGSLNRQPOK	136
QY	121	FPFALRGCPQSDSIAFLVDGSGSIIPHDFFRAKEFTSTVMEQLKSKTLFSLMQYSEEF	180
DB	137	FPFALRGCPQSDSIAFLVDGSGSIIPHDFFRAKEFTSTVMEQLKSKTLFSLMQYSEEF	196
QY	181	RIHFTFKFQPNPNSRLIKITQLLGHRTATGIRKVRLEFNITGARKNAFKILILI	240
DB	197	RIHFTFKFQPNPNSRLIKITQLLGHRTATGIRKVRLEFNITGARKNAFKILIVI	256
QY	241	TGDEKFGDPLGVEDVIPADREGVIRYVIGVDAPFRSEKSRQELNTVASKPRPDHVFQIN	300
DB	257	TGDEKFGDPLGVEDVIPADREGVIRYVIGVDAPFRSEKSRQELNTVASKPRPDHVFQVN	316

QY	301	NFEALKTIQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAITSGNPLLSVGSYDWAG	360
DB	317	NFEALKTIQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAITSGNPLLSVGSYDWAG	376
QY	361	GVFLYTSKEKSTFINWTRVDSMDNDAYLGAAAAIILNRVQSLVLGARYQHIGLVAMFR	420
DB	377	GVFLYTSKEKSTFINWTRVDSMDNDAYLGAAAAIILNRVQSLVLGARYQHIGLVAMFR	436
QY	421	QNTGHWESNANVKGTCIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETRGQGVSVCP	480
DB	437	QNTGHWESNANVKGTCIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETRGQGVSVCP	496
QY	481	PRGQARWQCDVAVLYGEGQPMGRFGAALTIVLGDVNGDKLTVAIGAPGEEDNRGAIVLF	540
DB	497	PRGQARWQCDVAVLYGEGQPMGRFGAALTIVLGDVNGDKLTVAIGAPGEEDNRGAIVLF	556
QY	541	HCTSGSGISPSHSQRIAGSKLSPRLQYFGQSSGGQDLTMDGLVLTVAQGHVLLLRSQ	600
DB	557	HCTSGSGISPSHSQRIAGSKLSPRLQYFGQSSGGQDLTMDGLVLTVAQGHVLLLRSQ	616
QY	601	PVLRVKALMEFNPREVARNVFECDQVYKVGKAGEVRVCLHVQKSTRDLRLRGGQIQSVVT	660
DB	617	PVLRVKALMEFNPREVARNVFECDQVYKVGKAGEVRVCLHVQKSTRDLRLRGGQIQSVVT	676
QY	661	YDLALDSGPHSRVAFNETKSTRQTVLGTTCETIKLQLPNCIEDPVSPVILRLNF	720
DB	677	YDLALDSGPHSRVAFNETKSTRQTVLGTTCETIKLQLPNCIEDPVSPVILRLNF	736
QY	721	SLVGTPLSFAFGNLRPVLAEDAQLFTALPFPFKNCGNDNICQDDLITFSPMSLDCLVVG	780
DB	737	SLVGTPLSFAFGNLRPVLAEDAQLFTALPFPFKNCGNDNICQDDLITFSPMSLDCLVVG	796
QY	781	GRPEFNVTVTRNDGSDSYRTQVTPPDLDSVRKYSTLONORSORSLRACESSASTEV	840
DB	797	GRPEFNVTVTRNDGSDSYRTQVTPPDLDSVRKYSTLONORSORSLRACESSASTEV	856
QY	841	SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF	900
DB	857	SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF	916
QY	901	QLELPKYAVYVMTSHGKYSTKYLNTASENTSRVWQHGYOVSNLQORSIPISLAVLPV	960
DB	917	QLELPKYAVYVMTSHGKYSTKYLNTASENTSRVWQHGYOVSNLQORSIPISLAVLPV	976
QY	961	RLMOTVIMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVAVNCIAVCORIQDIP	1020
DB	977	RLMOTVIMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVAVNCIAVCORIQDIP	1036
QY	1021	FFGIQBEFNATLKGSLSPDWYIKTSINHLIIVSTABILFNDSVFTLLPGQCAFVRSQTET	1080
DB	1037	FFGIQBEFNATLKGSLSPDWYIKTSINHLIIVSTABILFNDSVFTLLPGQCAFVRSQTET	1096
QY	1081	KVEPFEPVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMHSEGGPPGAEQ	1137
DB	1097	KVEPFEPVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMHSEGGPPGAEQ	1153

RESULT 2
S00551
leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
N;Alternate names: complement-3 receptor alpha chain
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 22-Oct-1999
C;Accession: S00551; 159078
R;Pytela, R.
EMBO J. 7, 1371-1378, 1988
A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the ;
A;Reference number: S00551; MUID:88312584; PMID:3044779
A;Accession: S00551
A;Molecule type: DNA
A;Residues: 1-1153 <PPT>
A;Cross-references: EMBL:X07640; NID:G52982; PIDN:CAA30479.1; PID:G52983
A;Note: the authors translated the codon CAC for residue 569 as Gin

R;Saetre, L.; Roman, J.M.; Teplov, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A;Title: A partial genomic DNA clone for the alpha subunit of the mouse complement receptor
A;Reference number: I59078; MUID:86287312; PMID:2942940
A;Accession: I59078
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 11-44 <RES>
A;Cross-references: GB:M14293; NID:9198993; PIDN:AAA39484.1; PID:9554193
C;Genetics:
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homodimer
C;Keywords: cell adhesion; glycoprotein; transmembrane protein
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 76.1%; Score 4470; DB 2; Length 1153;
Best Local Similarity 73.9%; Pred. No. 1.2e-302;
Matches 841; Conservative 146; Mismatches 149; Indels 2; Gaps 2;

Qy 1 FNLOTENAVTFQENARGFGQSVQVQGGSRVVVGAPQRIIVANQROSLVQCYSTGSCBPI 60
Db 17 FNLDTEHPMTFOENAKFGQNVVQLGTSVVVAAPQEAQVAVNQTGALYQCYSTSRCHPI 76

Qy 61 RLOVPVAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTVYKGLCFGLFGLNLRQPOK 120
Db 77 PLOVPEAVNMSLGLSLAVSTVPQOLLACGPTVHOTCKENTVYKGLCFGLNLRPPQ 136

Qy 121 FPEALRGCPQEDSDIAFLVDGSGSIIPEDFRRAKEFIISTVMBQKSKTLFSLMOYSBEP 180
Db 137 FPEALRGCPQEDSDIVFLDGGSSINIDFQMKSFVSTVMBQKSKTLFSLMOYSDEF 196

Qy 181 RHFTPEKQNPMPRLKIPITOLLGRTHATGIRKVRRELFNTNGARKNAKILILI 240
Db 197 RHFTFNDPKRPSRSHVSPKQLNGRTKTASGIRKVRRELFPHKNGARENAKILIVI 256

Qy 241 TDGEKFGDPLGVEDVPEADREGVIRYVIGVGDADFSEKSRQELNATVASKPPRDRVFPQIN 300
Db 257 TDGEKFGDPLGVEDVPEADREGVIRYVIGVGNFNAFKPQSRRELDTIASKPAGEHVFOVD 316

Qy 301 NFEALKTIONQREKIFAIEGTQGTSSSFHEKMQBGSAAITNSGILLSTVGSVDWAG 360
Db 317 NFEALNTIQNQEKEIFAIEGTQGTSSSFHEKMQBGSAAITNSGILLSTVGSVDWAG 376

Qy 361 GVFLYTSKSTFPMNTRVDSMDADYLGVAAILRNVRVQSLVLGAPYQHIGLVAMFR 420
Db 377 GAFLYTSKSTFPMNTRVDSMDADYLGVAAILRNVRVQSLVLGAPYQHIGLVAMFR 436

Qy 421 QNTGMWSEANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYETRGQGVSVCP 480
Db 437 ENPGTWEPHTSIKSGQISGYSFGASLCSVDMDADGNTNLLIGAPHYETRGQGVSVCP 496

Qy 481 PRGORARKWCDVLYGCEQOPGRCFAGALTVDVNGDKLTDAVIGAPGEEDNRGAVLF 540
Db 497 PRG-RARWQCEALHGDQGHQWGRFAGALTVDVNGDKLTDAVIGAPGEEDNRGAVLF 555

Qy 541 HGTSGSGISFESHQRIAGSKLSPLQYFGQSLGGQDLTMDGLDVLTVGAQGHVLLRSQ 600
Db 556 YGASIASLGSASHRIIAGHFSPLQYFGQSLGGQDLTMDGLDVLTVGAQGHVLLRSQ 615

Qy 601 PVLRVVAIMEFNPREVARNVFCNDQVVKGEAGEVRLVCHVOKSTRDLRREGIQSVVT 660
Db 616 PVLLEATMEFSPKKVARSVFAQEQVLRKNDAGEVRVLRVKNRNTKOLREGIDISTVT 675

Qy 661 YDLALSDGRPHSAVNETKNSRRQTVLGLTQCTETLKLQVFNCEIDPVPVILRLNF 720
Db 676 YDLALDVPVRSIRAFDETKNTNTRRTQVGLMKQKCTELKLILPDCVDSDVSPVILRLNY 735

Qy 721 SLVGTPLSAGNLRPVLAEDAQRLFTALPFFPKNCNDMNCQDDLSITFSFSLDCLVVG 780
Db 736 TLVGEPLRSFNLRPVLAEDAQRFPTAMPFFPKNCNDMNCQDDLSITFSFSLDCLVVG 795

Qy 781 GPREFNVTVVRNDGEDSVRTQVTFPPPLDLSYRKVSTLQNSORSWRL-ACESASSTSE 839
Db 796 GPQDFNVTVLRNDGEDSVRTQVTFPPPLDLSYRKVSTLQNSORSWRL-ACESASSTSE 855

Qy 840 VSGALKSTSCSINHPIFFPENSEVTFNITFDVDSKASLGNKLLKXANVTSENNMPRTNKT 899
Db 856 GHGALKSTTWNIHPIFFPENSEVTFNITFDVDSKASLGNKLLKXAVSENNMSRTHKTK 915

Qy 900 FOELPVPKYAVYVTVSHGVSTKYLNTASENTRVMOHQYQVSNLQORSIPISLVFLVP 959
Db 916 FOELPVPKYAYIMVTVSDSSIRYLNFTASEMSTKVQHQQYQFNNGQRLSPVSVVFWIP 975

Qy 960 VRLNQTIVDRPQVTPSENLSSSTCHTKERLPSHSDFLAELRKAPVAVNCISAVCORIQCDI 1019
Db 976 VQINNVTVWDRPQVTFISQNLSSCHTEQKSPHSNDRDLERTFVLNCSVAVCKRIQCDL 1035

Qy 1020 PFGIQIEFNATLKNLSFDWYIKTNSHLLIIVSTABILFNDSVFTLLPGGAFVRSQTE 1079
Db 1036 PSFNTQIEFNATLKNLSFDWYIKTNSHLLIIVSTABILFNDSVFTLLPGGAFVRSQTE 1095

Qy 1080 TKVEPPEVNPPLPIVGVSSVGGLLALLALITAAALYKLGFFKQYKDMKSEGGPPGCAEPQ 1137
Db 1096 TKVEPEVNPVPLPIVGVSSVGGLLALLALITAGLYKLGFFKQYKDMKSEGGPPGCAEPQ 1153

RESULT 3
RWHLIC
cell surface glycoprotein CD11c precursor - human
A;Alternate names: leukocyte adhesion receptor p150,95 alpha chain
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 22-Jun-1999
C;Accession: A36584; A35543; S00864
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 12750-12751, 1990
A;Reference number: A36584
A;Contents: erratum
A;Accession: A36584
A;Molecule type: DNA
A;Residues: 1-1163 <COR>
A;Note: this revision to the sequence from reference A35543 includes the carboxyl end
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A;Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.
A;Reference number: A35543; MUID:90153906; PMID:2303426
A;Accession: A35543
A;Molecule type: DNA
A;Residues: 1-834 <CO2>
A;Note: this sequence has been revised in reference A36584
R;Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A;Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A;Reference number: S00864; MUID:88166645; PMID:3327687
A;Accession: S00864
A;Molecule type: mRNA
A;Residues: 1-755, 'L', 757-1163 <CO3>
A;Cross-references: GB:M81695; EMBL:Y00093; NID:9487829; PIDN:AAA59180.1; PID:9487830
A;Note: part of this sequence was confirmed by protein sequencing
C;Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my
C;Genetics:
A;Gene: GDB:ITGAX; CD11C
A;Cross-references: GDB:119758; OMIM:151510
A;Map position: 16p11.2-16p11.2
A;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homodimer; tandem repeat;
C;Keywords: calcium; cell adhesion; signal sequence #status predicted <SIG>
F;1-19/Domain: signal sequence #status predicted <MAT>
F;20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>
F;20-1107/Domain: extracellular #status predicted <EXT>
F;149-319/Domain: von Willebrand factor type A repeat homology <VWA4>
F;1108-1133/Domain: transmembrane #status predicted <TM>
F;1134-1163/Domain: intracellular #status predicted <INT>
F;61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre

Query Match	59.28;	Score 3479;	DB 1;	Length 1163;
Best Local Similarity	61.18;	Pred. No. 1.3e-233;		
Matches	690;	Conservative 141;	Mismatches 292;	Indels 6;
Gaps	4;			
QY	1	FNLDTENAMTFQ--ENARGFGQSVVQLQGRVVVVGAPQBEIVAAVNRGSLYQCDYSTGSCBPI	60	
DB	20	FNLDTTELTAFRVDSAGFGSDVQYANVWVVGAPQKITAAVNRGSLYQCDYSTGSCBPI	79	
QY	61	RLQVPVAVNMSGLSLAATTSPPQLACGTFVHOTCSNTYVYKGLCFGLGSLNRQPOK	120	
DB	80	GLQVPPVAVNMSGLSLAATTSPPQLACGTFVHOTCSNTYVYKGLCFGLGSLNRQPOK	137	
QY	121	FPEALRCPOBDSIAFLVDGSGSIIIPHDFFRAKEPFTSTVMEQLKKSKTLFSLMYSSEB	180	
DB	138	LPVSRQSCPBQBDIVFLIDGSGSISSENATMWNFVRAVISQFORSTQFSLMYSSEB	197	
QY	181	RIHTFFEPONNPRSLRIPITOLGRTATGIRKVVRELTNTGAKNFKILILI	240	
DB	198	QTHFTPEPRTSNPLSLASVHQLQGTFTATAIVNVHRLPHASGARPRDRTKLIVI	257	
QY	241	TGKFGDELYEDIVEADREGVIRVVGDAFRSEKROELNTVASKPPRDHVFQIN	300	
DB	258	TGKFGDELYEDIVEADREGVIRVVGDAFRSEKROELNTVASKPPRDHVFQIN	317	
QY	301	NFEALKTIONLRKIPAIKGTQSGSSSEHEHMSQSGPAAITSNGPLISTVGSYDVG	360	
DB	318	DFDALKDIONLRKIPAIKGTQSGSSSEHEHMSQSGPAAITSNGPLISTVGSYDVG	377	
QY	361	GVPLVTSKSKSTFNTMTREVDMDNDAVLGAAAIILNRVQSLVIGAPRYCHILGVMER	420	
DB	378	GAFIYPPMKSPTFLNMQENVMDRDSVLGISTELALWKGVSQSLVIGAPRYCHILGVMER	437	
QY	421	QNTQWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGQVSCVPL	480	
DB	438	QVSRQWRMAEVTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGQVSCVPL	497	
QY	481	PRGQARWCDVLYGSGQPGWFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYLIF	540	
DB	498	PRGQARWCDVLYGSGQPGWFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYLIF	556	
QY	541	HGTSGSGISPSHSORISAGSLSPRLQYFGQSLGQDLTMDGLVDLVGAGQVHLLRSQ	600	
DB	557	HGVLGPSISPSHSORISAGSLSPRLQYFGQSLGQDLTMDGLVDLVGAGQVHLLRSQ	616	
QY	601	PVLVKALMEPNPREVARNVFCNDQVYKGEAGEVRLVHVSQSTDRLEBQIQSVVT	660	
DB	617	PVLVGVSMQPIPAEIPRSACEBQVSEBTLVQSNICLYIDKRSKNLIGSLDQSSVT	676	
QY	661	YDLALDGRPHSAFVFNKSTRTQVGLTQTCSTLKLQPNCHIEDPVSPITVLRNF	720	
DB	677	LDLALDGRPHSAFVFNKSTRTQVGLTQTCSTLKLQPNCHIEDPVSPITVLRNF	736	
QY	721	SLVGTPLSAFNLPRVLABDQRLFTALPPFKVCGNDNICQDDLSITFPMGLDCLVWG	780	
DB	737	TLVGKPLLAFLNRLPMLAADAQRYTTASLPFKVCGNDNICQDDLSITFPMGLDCLVWG	796	
QY	781	GPRESFNTVVRNDEGSDYRQVTFPPPLDLSYRKVSTFLQNRQSRWRLACRSASTEV	840	
DB	797	SNLEHAEVWVNDGSDYRQVTFPPPLDLSYRKVSTFLQNRQSRWRLACRSASTEV	854	
QY	841	SGALKSTSCSINHPFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNTEB	900	
DB	855	SGTWTSTSCRINHLIFRGAQITFLATFDVSPKAVLGLDRLLLTANVSSENTPRTSKTF	914	
QY	901	QLELPVXVYVMTVSHCVSTKYNLFAS-ENTSRVMOCHOVSNLQORSLSISLVLVP	959	
DB	915	QLELPVXVYVMTVSHCVSTKYNLFAS-ENTSRVMOCHOVSNLQORSLSISLVLVP	974	
QY	960	VRLNQTIWDRPQVTFGENSLSTCHTERLPFSDHDFLAELKAPVWNCISIAVCOICDI	1019	
DB	975	VELNQEAVMWDEVSHPNQPSLRCSSEKIPAPASDFLAHQKPNVLDCSIAGCLRPRCDV	1034	
QY	1020	PPFGIOBEFNAFLKGNLSFDVYIYKTSNHLIIVSTAILFNDVSTFILLPQGAQVRSQTE	1079	

Db	1035	PSFSVQGEELDFLKNLSFGWVRLQKKVSVVAEITFTDTSVYSQLPQGEAFWRAQTT	1094	
QY	1080	TKVSEFVFNPLPLIVGSGVGLLILALITAAALYKLGKFFKRYKMMSE	1128	
DB	1095	TVLEKYKVNPTPLIVGSGVGLLILALITAAALYKLGKFFKRYKMMSE	1143	
RESULT 4				
S03308				
cell surface glycoprotein CD11a precursor - human				
N;Alternate names: leukocyte adhesion glycoprotein LFA-1 alpha chain; leukocyte function				
C;Species: Homo sapiens (man)				
C;Date: 28-Feb-1990 #sequence revision 28-Feb-1990 #text_change 20-Aug-1999				
C;Accession: S03308; A47565; A48759; S36044				
R;Larson, R.S.; Corbi, A.L.; Berman, L.; Springer, T.				
J. Cell Biol. 108, 703-712, 1989				
A;Title: Primary structure of the leukocyte function-associated molecule-1 alpha subunit				
A;Reference number: S03308; MUID:89139587; PMID:2537322				
A;Accession: S03308				
A;Molecule type: mRNA				
A;Residues: 1-1170 <LAB>				
A;Cross-references: EMBL:X00796; NID:931421; PIDN:CAA68747.1; PID:931422				
A;Note: part of this sequence was confirmed by protein sequencing				
R;Cornwell, R.D.; Gollan, K.A.; Hickstein, D.D.				
Proc. Natl. Acad. Sci. U.S.A. 90, 4221-4225, 1993				
A;Title: Description of the leukocyte function-associated antigen 1 (LFA-1 or CD11a) pr				
A;Reference number: A47565; MUID:93281759; PMID:8097887				
A;Accession: A47565				
A;Molecule type: DNA				
A;Residues: 1-20 <COR>				
A;Note: sequence extracted from NCBI backbone (NCBIN:130862, NCBIP:130863)				
R;Shelley, C.S.; Farokhzad, O.C.; Arnaout, M.A.				
Proc. Natl. Acad. Sci. U.S.A. 90, 5364-5368, 1993				
A;Title: Identification of cell-specific and developmentally regulated nuclear factors				
A;Reference number: A47565; MUID:93281759; PMID:8099450				
A;Accession: A47565				
A;Molecule type: DNA				
A;Residues: 1-20 <SHE>				
A;Cross-references: GB:M95609				
R;Nueda, A.; Lopez-Cabrera, M.; Vara, A.; Corbi, A.L.				
J. Biol. Chem. 268, 19305-19311, 1993				
A;Title: Characterization of the CD11a (alphaL, LFA-1alpha) integrin gene promoter.				
A;Reference number: A48759; MUID:93374910; PMID:8103515				
A;Accession: A48759				
A;Molecule type: DNA				
A;Residues: 1-20 <NUE>				
A;Cross-references: EMBL:222804; NID:9311405; PIDN:CAA80461.1; PID:9311406				
C;Genetics:				
A;Gene: GDB:ITGAL; CD11A				
A;Cross-references: GDB:119757; OMIM:153370				
A;Map position: 16p11.2-16p11.2				
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom				
C;Keywords: cell adhesion; cytoskeleton; glycoprotein; heterodimer; surface antigen; tri				
F;1-25/Domain: signal sequence #status predicted <Sig>				
F;26-1170/Product: leukocyte adhesion glycoprotein LFA-1 alpha chain #status predicted				
F;154-317/Domain: von Willebrand factor type A repeat homology <VWA2>				
Query Match 26.4%; Score 1553.5; DB 2; Length 1170;				
Best Local Similarity 34.7%; Pred. No. 1.7e-99;				
Matches 407; Conservative 206; Mismatches 460; Indels 101; Gaps 37;				
QY	1	FNLDTENAMTFQ--ENARGFGQSVVQLQGRVVVVGAPQBEIVAAVNRGSLYQCDYSTGSCB	59	
DB	26	YNLDVRGASFSPPRAGRHFGRVQLV-GNGVIVGAPGE--GNSTGSLYQCSGTGCHL	81	
QY	59	FIRLQVPVAVNMSGLSLAATTSPPQLACGTFVHOTCSNTYVYKGLCFGLGSLNR---	115	
DB	82	PVTLR-GSNYSKYLGWLTATDFTGSLILACDPGLSRTCDQNTYISGLCYLFRNLQGM	140	
QY	116	-QOQKFFBALRGCPQEDSDIAFLVDGSGSIIIPHDFFRAKEPFTSTVMEQLKKSKTLFSLM	174	
DB	141	LQGRPGQBCING---NVDLVFLDGSMSLPDEFQKILDFMKDVMKKLSNNTSYQFAAV	196	

175 QY SEEFRIHETKEPQNNPNSRLIKPIQTQALGRTHATGIRKVVRELFNITNGARKNAP 234
 197 QY STSTKTFPDSYVYKWDQDALLKXVHMLLNTFGAINVATEVREBELGARPDA 256
 235 QY KILILITDGEKFDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTVASKPPED 294
 257 QY KVLIIITDGE--ATDSGNIDAAKD---IIRVIGIKGHFKQTSQETLHKFASKPASE 309
 295 QY HYPQINNPEALKTIONQLBEKIPAEIGTQSGSSSEHEMSQEGSAAITNSGPLLSTVG 354
 310 QY FKILITDPEKLDLFTLOKQIYVIGTSKQDLTSFNMLSSSGISADLSRSHGVAVGAG 369
 355 QY SYDWAAGGVF-LYTSKEKSTFINNTRVDSMDNAYLQYAAA-IILNRVQSLVLGAPRYOH 412
 370 QY AXDWAGGFLDKADLQDDTFIGNELTPEVRAGYLGTVTWPSPKQTSLLASGAPRYOH 429
 413 QY IGLVAMFR--QNTGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEOT 470
 430 QY MGRVLLFQEPQGGHWSQVTHGTQISYFGSELGVDVDQDGETELLIGAPLYGSG 489
 471 QY RGGQSVCPPLPRQARWOCDAV--LYGEOGQFPGFRGAALTVLGVDNGDKLTDLVAIGAP 528
 490 QY RGRVFIY-----QRQLGFEEVSEHQDPGYPLGRFGEAITALTDINGDLVDVAVGAP 544
 529 QY GEDNRGAVYLFHGTSGSGISPSHSORIASGLSPRLQYFGQSLSGGQDLTWGGLVDLTV 588
 545 QY LEE--QCAVYVFNHGRG-GLSPQSPQIRGTQVLSGIOWFGRSHGVKDLGGLADVAV 601
 589 QY GAQGHVLLRSQPLVAKVAMFNPPEVARNVPCNDQVY-KKKGAGEVRCVCLHVKQSR 647
 602 QY GAGSQMIVLSRPVDMVIMLSPFPAEIPVHEVECSYSTSNKRGVNTICQI-KSLY 660
 648 QY DLREQIOQSVVYTDIALDSGRPHSAVFNENKSTRQVQLGHTQTCETLKLQPLNCI 707
 661 QY PQP-QGRVANLTYTQLDGHRTTRRGLPFGGRHLLRNIAVT-TSMSCDPSFHPVCV 718
 708 QY EEPVSPVILRLNESL---VGTPLS--AFGN-----LRPVLAEDAQLFTALPEFKNCGN 757
 719 QY QDLISPINVSLNPSLMEERGGTRDQSAQKQIPILLRSLHSETWEI-----PFENKCGE 773
 758 QY DNIQDDLSITPSFMSLDCVLVGGPREFNVTVTRNDGDSYRTQVTPFPPLDLRSKYVS 817
 774 QY DKKEANLVSPSPARSRALRUTAPASLSVELSLNLEDAVWQLDLHFPPLSFRKVE 833
 818 QY TLONQORSWRLACES--ASSTEVSGALKSTSCSINHIPPENSEVTNITFDVDSKAS 875
 834 QY ML---KPHSQIIPVSCBELPEESRLLSRAL---SCNVSSPIFKAGHSVALQMMFNTLVNS 887
 876 QY LGNKLLKLVNYSENN---MPTNKTETQLPLPVKYAVYVTVSHGVSTKYLNTFASGN 931
 888 QY WEDSVELHANYTCNEDSDLEDNSATTI---IPILYPINILIQDQEDSTLYVSTPKGP 944
 932 QY TSVVMQHQYOV---SNLQORSIP-LSLFLVPLVRLNQTVINDRPQVTFSENLSSTCHTK- 986
 945 QY KIHQVKHMTQVRIQPSIHDNITPLEAVGVQPPSEGPITHQWSVQMEPPV--PCHVED 1002
 987 QY -RLPSHSD--FLAELKAPVNCISAVCQRQCDIPFPGIIEEENATLKNLSFDWIK 1043
 1003 QY LERLPDAAEPCPLGALFRCPV-----PFQELVQVIGLLELVGEIE 1044
 1044 QY TSHNHLIVSTABILFNDVFTLLPQGGAFVRSQETETKVEPEVENPFLIYVGGSVGGL 1103
 1045 QY AS-SMPSLCSLSISNSKHFHLYGSNALS-AQVVMKVDVYVYKQMLYLYVLSIGIGLL 1102
 1104 QY LLALITAAIYKLGFPKQYKDMSEG-GPPGAP 1136
 1103 QY LLLLPIVLYKVGFKRNLKEKMEAGRGVFNIGP 1136

RESULT 5
 I56126
 lymphocyte fuction-associated molecule-1-alpha - mouse

C:Species: Mus musculus (house mouse)
 C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Aug-1999
 C:Accession: I56126
 R:Kaufmann, Y.; Tseng, E.; Springer, T.A.
 J. Immunol. 147, 369-374, 1991
 A:Title: Cloning of the murine lymphocyte function-associated molecule-1 alpha-subunit
 A:Reference number: I56126; MUID:91268576; PMID:2051027
 A:Accession: I56126
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1163 <RBS>
 A:Cross-references: GB:M60778; NID:g198785; PIDN:AAA39426.1; PID:g198786
 C:Gene: LFA-1
 C:Genetics:
 F:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
 F:151-315/Domain: von Willebrand factor type A repeat homology <VMA>

Query Match 26.2%; Score 1538.5; DB 2; Length 1163;
 Best Local Similarity 34.2%; Pred. No. 1.9e-98;
 Matches 401; Conservative 214; Mismatches 458; Indels 99; Gaps 36;

QY 1 FNLDTENANTFOENA-RGFGQSVVQLQGSRRVVGVAPOBIVAANORGSLYQCDYSTGSCPEP 59
 DB 24 YNLDTPTQSFQAQAGRHFGYQVLIQEDG-VVVGARGE--GDNTGGLYHCTISSEFPQP 79
 QY 60 IRLQVPEAVNMSLGLSLAATTSPPQLACQPTVHQTCSNTYVYKGLCFLFGSNLRQOPQ 119
 DB 80 VSLH-GSNHTSKYLMTLATDAKGLSLACDPEGLSRTCDQNTYLSGLCYLPQSLGPM 138
 QY 120 KPPALRCQPCQDSDIAFLVDGSGSIIIPHEPRAKERISTVMEOLKSKTKTFLSMQYSEE 179
 DB 139 QNRFAVQCMKGVLDLFLPDGQSLDRKQPEKILEFMKQVWRKLSNTSYQFAAVQFSTD 198
 QY 180 FRIHFTPEKF-QNNFNPRSLIKPIQTQLGRTHATGIRKVVRELFNITNGARKNAFKILI 238
 DB 199 CRTEPTFLDYVQKNKPNPDLVLSGVPMLFELTFRAINVYVAHVFKBBSGARPDATKVLV 258
 QY 239 LITDGEKFDPLGYEDVTPHADREG-----VIRYVIGVGDAPFRSEKSRQELNTVASKP 291
 DB 259 IITDG-----EASDKGNISAADHTIYRIIGIKGHFVSVQKQTLHIFASEP 304
 QY 292 PRDHVQINNPEALKTIQNLREKIFAIEGTQTSSSSEHEMSQEGSAAITNSGPLL 351
 DB 305 VEEFVKILDTLTKLDLFTDQIRRIYALIEGNNRQDLTSFNMLSSSGISADLSKHAHV 364
 QY 352 TVGSDVWAGGVF-LYTSKEKSTFINNTRVDSMDNAYLQYAAA-IILNRVQSLVLGAPR 409
 DB 365 AVGAKDWAGGFLDLREDLQCATFVQBEPLTSDVREGYLGYTVAMWTSRSSPLLAAGAP 424
 QY 410 YOHIGLVAMFR--QNTGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHY 467
 DB 425 YQHVQVLLFOAPAGGEWNTQKIEGTQISYFGSELCSVDLQDGEAEILLIGAPLP 484
 QY 468 EOTRGQSVCPPLPRQARWOCDAVLYGEOGQFPGFRGAALTVLGVDNGDKLTDVAIGA 527
 DB 485 GEORGRVFTY---QRRQSLPEMVSELQGDGPGYPLGRFGAALTALTDINGDLTDVAIGA 541
 QY 528 PGEEDNCAVYLFHGTSGSGISPSHSORIASGLSPRLQYFGQSLSGGQDLTWGGLVDLT 587
 DB 542 PLEB--QCAVYVFNHGRG-GLSPQSPQIRGTQVLSGIOWFGRSHGVKDLGGLADVA 598
 QY 588 VCAQGHVLLRSQPLVAKVAMFNPPEVARNVPCNDQVYKKEAG-EVRVCLHVKQST 646
 DB 599 VCAEGRVVVLSRRPVDVWVTELSFSPETIPVHEVECSYSAREEQKHGKVLKACFRPKLT 658
 QY 647 RDLREBQIQSVVYTDIALDSGRPHSAVFNENKSTRQVQLGHTQTCETLKLQPLNC 706
 DB 659 PQ--FOGRLANLSTYQLDGHMRMRSLFPGDGHSELSGNTSITTP-DKSLDFFHFFPIC 715
 QY 707 IEDPVSPVILRLNESL---VGTPLSAGFN-LRPVLAEDAQLFTALPEFKNCGNDITCQ 762
 DB 716 IQDLISPINVSLNPSLMEERGGTRDQSAQKQIPILLRPSIHTV-TKEIPFENKCGEDKCE 774

QY	763	DLSTITFSMSLDCLVGGP-----REFNVTVTVVRDGEDSYRTQVTFPPFLDLSTYRKV	816
DB	775	ANLTLSSPARS-----GFLRLMSSASLAVETWLSNGEDAYVVRLLDLDFPRGLSFRKV	827
QY	817	STLQNRQSRWELACESASTEVSCAL-KSTSCSINHPIFPENSEVTFNITFVDVSKAS	875
DB	828	EMLQ-----PHSRMPVSCBEL--TEGSLTKTKLVKNVSSPIFKAGQEVSLQVNFILNSS	882
QY	876	LGNKLLKANVTSEN-NMPTNKTEPQLBELPKVYAYVMVTVSHGVSTKYLTANTASNTSR	934
DB	883	WEDFVELNGTVHCENENSSLOEDNSAATHIPVLPVNVILTKQENSTLIYFTPKGPKTQ	942
QY	935	VMOHOYQVNSLGNORSIPISLVLVPLVRLNQTVWRPQ-----VTPSENLS-----TCHTK	986
DB	943	QVOHVTVQ-----RIQPSAYDENMPT-LEALGVPRPHSEDLITITWSTQVDFLVTCSE	996
QY	987	E-RLPESHDFLAEALRKAPVNVCSIAVCRIQCDIPFPGIQEENFATLKNLSFDMYIKTS	1045
DB	997	DLKRPSE--AEQPCLPGV-----QFRCPIVE--RWEILIOVTVGVLSKEIKAS	1042
QY	1046	HNHLIVSTAELFNDVSVTLIPQCAFVRSQTETKVPPEVNPVPLPLVGVSSVGLLILL	1105
DB	1043	-STLSLCSLSVSFNSSKHFHLYGSKA-SEAQVVLVKVLIHKEMLHVVYLSGIGELVILL	1100
QY	1106	ALITAAALYKLGFPKROVKDM-SEGPPPGAEP	1136
DB	1101	FLIFLALYKVGFPKRLKEKMEADGCVNGSP	1132
RESULT 6			
A53213			
Integrin alpha-E chain - human			
C:Species: Homo sapiens (man)			
C>Date: 19-Oct-1995 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001			
C/Accession: A53213			
R/Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.			
J. Biol. Chem. 269, 6016-6025, 1994			
A>Title: Molecular cloning of the human mucosal lymphocyte integrin alpha (E) subunit. Un			
A/Reference number: A53213; PMID:94164962; PMID:8119947			
A/Accession: A53213			
A>Status: preliminary			
A/Molecule type: mRNA			
A/Residues: 1-1179 <SHA>			
A/Cross-references: GB:L25851; NID:9457244; PID:9457245			
C/Genetics:			
A/Gene: ITGAE			
A/Cross-references: GDB:330801			
A/Map position: 17p13			
A/Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom			
F;199-371/Domain: von Willebrand factor type A repeat homology <VWA3>			
Query Match 19.6%; Score 1153; DB 2; Length 1179;			
Best Local Similarity 29.1%; Pred. No. 1.3e-71;			
Matches 342; Conservative 214; Mismatches 449; Indels 172; Gaps 39;			
QY	45	GSLYQCDYSTGS--CEPI-RLQVP-----VEAVNMSLGLSLAATTSPPQLLACGPTVHQ	95
DB	65	GPLHRCSLVQDEILCHPVEHVPIPKGRHGVTVVRSRHGVLCI-----QVLVRP--HS	117
QY	96	TCSENTVVKGLCFLFGSNLRQPPQ-----	119
DB	118	LSSELT---GTCSLLGPDLRPQAQNFDELLENLDPDARVDTGDCVSNKEGGEDDVNTA	174
QY	120	KPEALRGCPQED-----SDIAPLVDSGSIIPHDFRAKEPISTVMEQL--	164
DB	175	RQRALKEKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEED	234
QY	165	KSKTKLPSLMOYSEERFIHFTFKFQNNFNPRSLIKPIQTQLGRTHTATGIRKVVRELFN	224
DB	235	KCEFCNFALVOYGVQVITQTEFDRSDQVNASLARVQNIQVGSVTKTASAMQHVLDSIFT	294
QY	225	ITNGAKRNAPKILILITDGEKFGDPLCYEDVPEADREGVIRVIVGVGDAFSEKSRQL	284

DB	295	SSHGRSKAKVMVVLTDGGIFEDFLNLTTVINSPKMGVERPAIGVGEERPKSARTAREL	354
QY	285	NTVASKPPDRHVQINNPEALKITIONOLREKIFAIEGTQTGSSSPHEHMQEFSAAIT	344
DB	355	NLIASDPDETHAFKVTVMALDGLLSKLVYIISMEGT---VGDALHYQLAQIGFSAQIL	411
QY	345	SNGP-LLSTGVSVDWAGSVFLY-TSKEKSTPINNTRVDSMDMDA---YLGYAAAILLRN	398
DB	412	DERQVLGAVGRFVWGGALLYDTRSRGRFLNQTAARAADEAAQYSYLGAVAVLHKT	471
QY	399	RVQSLVLGAPRYQHIGLVAMFR-QNTGWMESNAV-KGTQIGAYFGASLCSVDVDSNGST	456
DB	472	CSLSVAGAPQVKKHG--AVPELQKEGRASFLPVEGEQMGSYFSGELCEPVDIMDGGT	529
QY	457	DVLVLGAPHYVEQTFGGQSVCPPLPRGQRAWQCDVLYGQGGQWGRFGAALTVLGDNV	516
DB	530	DFLVLAAPFYHVHGEGRVYVRLSB-ODGSFSLARILSHGPGTNRARFGFAMAMGDL	588
QY	517	GDKLTDAIGAIP-----GEEDNR--GAVLPHGTSGSGISPSHSQRIAGSKLSPLRQYFGQ	570
DB	599	QDKLTDAIGAIPLEGFGADDCASFGSVIYNG-HWDGLSASPSQIRIRASTVAPGLQYFGM	647
QY	571	SLSGGQDVTMDGLVDLTVGAQGHVLLRSQPLVAKALMEFNPREVARNVTECDNVYKG	630
DB	648	SMAGGFD:SGDGLADITVTGLGQAVVPRSRPVRLKVSMAFTPSALP-----LGP	697
QY	631	KEAGEVRVCLHVQKSTRDLREGQIQSVVTVYDLALDSGRPHSRVAFNETKXSTRQTQVL	690
DB	698	NGVNVRLCFEI-SSVTTASESGLREALNFTLDVGVKQRRRLQCSVDVRSCLGCLREWS	756
QY	691	GLTQTCETLKLQLEN-----CHIEDPVSPVLRLNLSVGTPLSAGFNLRPLVLAEDARLFT	746
DB	757	SGSQCEDL-LMPTEGELCEEDCFNSASRVSYQL-QTPEGQTDHPQPIIDRYTEPAI	814
QY	747	ALFPFKKNCNDNICODDLSITTFSMELDCLVWGPREFNVTVTVRNDGDSYRTQVTF	806
DB	815	FQLPYKAKCKKLFCVABLQIA-TTVSQEELVGLTKELTLNINLTNSGEDSYMTSMALN	873
QY	807	FPDLISYKAVSTLQNRQSRWRACSSASTEVSGALKSTSCSINHPIFPENSEVTENI	866
DB	874	YPRNLQ-----LKEQKPPSPNIQCDPPQV--ASVLIIMNCRIGHFVL-KRSSAHVSV	923
QY	867	TFDVSQKASLGNKULLKANVTSENN---MPRTNKTEFO---LELVKYAVVMVTVSHGV	919
DB	924	VQQLSENAFPNRTADITVTVNSNERSLANETHLQFRHGFVAVLSKPSIMYVNTQGL	983
QY	920	S--TKYLNFTASENTSRVMQHYQVSNLQSRSLPISLVLPVRLNQTVIWDROPVTFSE	977
DB	984	SHHKEFLPHVHGEN---LFGAEYQ-----LQICVPTKLRGLQVAAVKCLTRTQ	1028
QY	978	NLSSTCHTKERLPKSHSDFLAELRKAPVNVCSIAVCRIQCDIPFPGIOSEBNATLKGNSL	1037
DB	1029	ASTVCTWSQERACAYSS-VQHVEMHVSVCVIA-----SDKENVTVAABIS	1073
QY	1038	PDWYIKTSHNHLIIVST-----AEILFNDSVFTLLPQGGAFVRSQTKETKVEPPEVNP	1091
DB	1074	WD-----HSELLKQVTELQILGEISFNKSLYELNAENH-RTKITVWFLKDEKXHS	1125
QY	1092	PLTVGSSVGGILLALITAAALYKLGFPKROVKDMSE	1128
DB	1126	PIIINGSVGGLLVILVILVILFKGFPKRYQOOLNLE	1162

RESULT 7.

A5226
Integrin alpha-1 chain - human (fragment)

C:Species: Homo sapiens (man)
C>Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 15-Sep-2003
C/Accession: A45226
R/Briesewitz, R.; Epstein, M.R.; Marcantonio, E.E.
J. Biol. Chem. 268, 2989-2996, 1993
A>Title: Expression of native and truncated forms of the human integrin alpha 1 subunit
A/Reference number: A45226; PMID:93155124; PMID:8428973

392 AAILNRVQSLVGLAPRYOHIGLVAMFRONTGMWESNANV-----KGTQIGAYFGASL 445
Db VASISGNSVHFVAGAPRANYTQIVLYSVN-----ENGWTVIQORGEQIGSIFGSVL 484
446 CSVDVDSNGSTDLVLGAPHYBQTR--GGQSVVCPPLRQGRARWQCDVAVLYGEOQPGW 503
Db CAVDVNDKDTIDVLLVAGAPWMDLKKRGGVYLFITKG--ILNMH--OFLEGPNGLENA 541
450 RFGAALTVGLDVGNDKLTDAVAGPGEEDNRGAVLPHGTSGSGISPSHSORAGS--KL 561
Db RFGSAALASDIINDGFENDVIVGSPLENQNSGAVIYNGHEG--ILRYSQKILASDRAP 600
452 SPRLQYFGSGSLGQBLTMDGLVDLVGAGQHVLLRSQFVLKRVKAIMFNFNPREVARNVP 621
Db SSHLQYFGRSLDGYGLNGDSITDVSAGAFQVQVQLWSQSIADSVSDAFTPKKI--TL 658
454 ECNDQVWKGAEGERVCLHVQKSTRDLREGQIOSVVTYDLALD-----SGRPHSAVEN 677
Db NRVAEI-----KLKLP-----SAKPRPTNQNNQVAIVTITIDBDQSSRVISGLPK 707
456 ETKNSTRTQVGLGLTQCE--TLKQLPNCIEDPVPSPVLRINFLS--VGTPLSAFCNL 733
Db ENNERCLQKTMIVSQORCEYIIHQEPS--DIISPLNLCWNISLENGT-----756
458 RVLADDAQRLFTALPFPFKKNCNNDNICODDLSITP-----SFMSLDCLVVGPRENVT 789
Db NPALAEYSETVYKVPFPHKXGDDGVCISDLVNVQQLPATQQQPFIVSNQKRLTFSV 816
460 TVENDGEDSVRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASST--EYSGALKSTS 848
Db OLKXKESAVNTBIIVDPSENLEF-----ASHMPVDGTEVTQIASSQKSVT 864
462 CSINHPFPENSEVTNITFDVDSKASLGNKLLKANKVTSNNMPTNTKTSQLELPVKY 908
Db CNVGPALKSKQVTFITNDFNLQ--NLQQAISIFRALSESQENNAADNSVNLKLSLY 923
464 AVTVVTSHGVSSTKYNLFTASENTSRVMOHQVSNLQOR-----SLPISLVPLV 958
Db DASIHIT--RSTNINFEVSLDGNVSVV--HSFE--DIGKPIKIVTGVSPVSWA---976
466 PVRLNQTVIWDREPQVTPSEN---LSSTCHTKB-----RLPSHSDFLAE- 998
Db SVIHIPOYTKXNPLMYLTVGHTDQAGDISCBAEINPLKIGQTSSSVSFKSEN 1030
468 LRKAPVNCIAVCQIQCDIPFGIOEBEENATLXGNLFDWIKYSHNELLIVSTAEI- 1057
Db FRHKLNCRTASCNIMCWLRLQVKGEYFLNVSTRIMNGTFAASTFQTVQVLTAAEID 1090
470 LFNDVSFTL-----LPGQAFVRSOTETKVBEPFE--VPNPLPLIVGSSVGGLLALLITA 1110
Db TYNPQIVIEENTVTP-----LITMKPHEKVEVPTGVIGSVIAGILLALLA 1140
472 ALYKLGFFPKQYKDM 1125
Db ILWKLGFPRKRYKEM 1155

RESULT 9
S44142
VLA-2 protein homolog - mouse
C/Species: Mus musculus (house mouse)
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Sep-2003
C/Accession: S44142
R/Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze
submitted to the EMBL Data Library, January 1994
A/Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not
A/Reference number: S44142
A/Accession: S44142
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-1178 <EDE>
A/Cross-references: EMBL:Z29987; NID:q473098; PIDN:CAA82877.1; PID:q473099
P:169-344/Domain: von Willebrand factor type A repeat homology <WAA2>

Query Match 18.2%; Score 1071; DB 2; Length 1178;
Best Local Similarity 27.9%; Pred. No. 6.9e-66;
Matches 342; Conservative 208; Mismatches 487; Indels 188; Gaps 44;
QY 1 FNLDTENAMTQ--ENARGFGOSVVLQ---QSRVVVWVGAQBEIVAAQNGSGIYQC--DYST 54
Db YNVGLPGAKIPSGPSSEQGYVQQLTNPQGNWLLVGSFSPFNRMGDVYKCPVDLPT 86
55 GSCPIRLQ-----VPVBAVNSLGLSLAATTPPQLLACGPTVHQTCSNTYVKGJL 107
Db ATCEKLNINQNSASISNTVETIKTNMSLGLTLTRNPTGGTGLCGPLWAHQCGNOYATGIC 146
108 FLFGSLNLFQQPQ---KFPALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKXPFISTVMQOL 164
Db ---SDVSPDFQRLTSPSPAVQACPSL--VDVVVWCDSENSTYV--WEAVQNFVLFVFTGL 199
165 K--KSKTLFSLMVSSEPRHIFTKEPQNPNPRLSLIKPITQLLG--RHTATGIRKVVRE 221
200 DIGPKKTOVALIQVANEPRIIIFNLNDFETKEDMVOATSETQHGDLTNTPFAIEPARDY 259
222 LFNITNGARKNAFKILITDGEKFGDPLGYEDVPEADREGVIRYVIGV-----GDAPR 276
260 AYSOTSGRPGATKVMVVVTDGESH--DGSKLKTVIQCCNDDEILRPGIAVLGYLNRNALD 318
277 SEKSRQELNVTASPPRDRHVFQINNFEALKTIONQLRKKIPIABGTQTGSSSFEHEMSQ 336
319 TKNLKIKALASTPTERYFFNVADAEALLEKAGTLGEQIFSIETGVQG--GDNQPMENAO 377
337 EGFS--AITNGPILLSVGSYDWAGGVFLVTSKSTFINMT--RVDSDMN--DAYLGYA 391
378 VGFSADYAPQDILMLGAVGAFDMSGTLVQETSHKPVIFPKQAPQVQLQDRNHSFLQYS 437
392 AAILNRNVQSLVGLAPRYOHIGLVAMFRONTGMWESNANV-----KGTQIGAYFGASLCS 447
438 VAAISTEDGVHFAVAGAPRANTTQIVLYSVNK--QGNVTVIQSHRGDQIGSYFGSLCS 494
448 VVDVDSNGSTDLVLGAPHYBQTR--GGQSVVCPPLRQGRARWQCDVAVLYGEOQPGWRF 505
495 VVDVDDTITDVLVAGAPYMDLKKRGGVYLFITKGILNCHQ---FLEGEPTGNARP 551
506 GAALTVGLDVGNDKLTDAVAGPGEEDNRGAVLPHGTSGSGISPSHSORAGSLSR- 564
552 GSAIAALSDINDGENDVIVGSPVENENSGAVIYNGHQGT--IRTKYSQKILSGNAPRR 610
565 -LQYEGOSLGGODLTMDGLVDLVGAGQHVLLRSQFVLKRVKAIMFNFNPREVARNVPEC 623
611 HLQFFGRSLDGYGLNGDSITDVSIGALQVQLWSQSIADVIAEALFTP-----660
624 NDQVWVKGAEGERVCLHVQKSTRDLREGQIOSVVTYDLALD---SCRPHSAVENET 679
661 -DKITLLNKDAKITLCLCFRAEFPAGQNNQV--AILFNMTILDAGHSSRVTSRGVFRFN 717
680 KNSRTRQTVLGLTQTCET--LKLQLPNCIEDPVPSPVLRINFLNFSLVGTPLSAFGNLRPVL 737
718 SERPLQKNVWNEVQKSEHHISIQPS---DVVNPLDLRVDISLENPGTS-----PAL 768
738 ABDAQRLFTALPFPFKKNCNNDNICODDLSI-----TFSPMSLDCLVVGPPREFNVT 788
769 EAYSETVYKVPFIPFYKCEGSDGICISDLILDVQQLPATQTSF-----IVSNQKRLTFS 823
789 VTVRNDGEDSVRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASST--EYSGALKST 847
824 VILKRGESAVNTVVLAEFSENLEF-----ASFNPMPVDGTEVTCEVGSQSKSV 871
848 SCINHPFPENSEVTNITFDVDSKASLGNKLLKANKVTSNNMPTNTKTB--FQLELP 905
872 TCDVGYPAKSEQQVTFITNDFNLQ--NLQQAIFNFOAPSESQ--ETNKADNSVSLTIP 928
906 VKYAVVWVTSHGVSSTKYNLFTASENTSRVMOHQVSNLQOR-----SLPISLV 955
929 ILYDAELMLT--RSTNINFEYIISDENAPSVIK---SVEDIGPKFIFSLKVTAGSAPVSWA 984

QY 956 FLV-----PVRINQIVWDRQVTF-SEVLX 980
D 985 LVTHIPQYTKENKPLLYLTGIOTDAGDISCTAEINPLKLPHTA-----PSVSFKNENFR 1040
QY 981 STCHKERLPSSHDFLAELRKAPVNGSLAVCQRIQCIDIPFGQIEBENATLKNLSFDW 1040
D 1041 ---HTKE-----LDCEITSCSNITCWLKDLHMKAEYFNVTRVWRT 1086
QY 1041 YIKSHNHLIVSTAEILFNDVSTLLPGQGAFFVRSOTETKVRPFVNPPLPLIVGSSVG 1100
D 1081 FAASTFTQVLTAAAEIDTHNPQLFVIEENAVTILMIMKTEKAEVPT--GVIIGSIIA 1138
QY 1101 GILLALLITAAALYKLGPPKROKDM 1125
D 1139 GILLALLITAAALYKLGPPKROKDM 1163

RESULT 10
A33998
integrin alpha-2 chain precursor - human
N:Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 chain
C:Species: Homo sapiens (man)
C:Date: 30-Mar-1990 #sequence revision 18-Sep-1992 #text_change 15-Sep-2003
C:Accession: A33998; B56793; A53117
R:Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A:Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet)
A:Reference number: A33998; MUID:89309879; PMID:2545729
A:Accession: A33998
A:Molecule type: mRNA
A:Residues: 1-1181 <TAK>
A:Cross-references: GB:X17033; NID:G33906; PIDN:CAA34894.1; PID:G33907
A:Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
B:Catimel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A:Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIIb and GPIIc)
A:Reference number: A56793; MUID:92061944; PMID:1953640
A:Accession: B56793
A:Molecule type: protein
A:Residues: 30-43 <CAT>
A:Experimental source: Platelet
R:Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A:Title: The human alpha-2 integrin gene promoter. Identification of positive and negative regulatory elements
A:Reference number: A53117; MUID:94103255; PMID:8276836
A:Accession: A53117
A:Molecule type: DNA
A:Residues: 1-16, v. 18-21 <ZUT>
A:Cross-references: GB:L24121; NID:G400342; PIDN:AAA16619.2; PID:G4583535
A:Note: authors translated the codon GAT for residue 17 as Leu
C:Genetics:
A:Gene: GDB:ITCA2; CD49B
A:Cross-references: GDB:128031; OMIM:192974
A:Map position: 5q11.1-5q11.2
C:Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-1133/Domain: extracellular #status predicted <EXT>
F:172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1134-1154/Domain: transmembrane #status predicted <TM>
F:1155-1181/Domain: intracellular #status predicted <CYT>
F:105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 18.2%; Score 1068; DB 2; Length 1181;
Best Local Similarity 26.9%; Pred. No. 1.1e-65;
Matches 329; Conservative 214; Mismatches 494; Indels 188; Gaps 43;

QY 1 FNLDTENAMTFQ-ENARFGQSWQL---QGSRRVVGAPQIRVAANQRGSLYQC--DYST 54
D 30 YNVGLPEAKIFGSPSEQFGYAVQOFINPKGNLLVSGSPWSPENRNGDVVKCFVDLST 89
QY 55 GSCPEIRLO-----VPVEAVNMLGLSLAATTSFPOLLACGPTVHQTCSNTYKGLC 107
D 90 ATCEKINLQSTSTSIPIVNTETKNTSLGLILTNMGTGGLTCGPLWAOCCNQYITTCVC 149

QY 108 FLGSLNRQOPKPPPALRGCPQSDSDIAFLVDGSGSIIPHDFRRAKEFIISTVMEQLK-- 165
D 150 SDISDPF-QLSASFSPATPCPSL-IDVVVVCDSNSIYPWD--AVNFKLEKFFVQGLDIG 205
QY 166 KSXTLSLAQYSSEFRIHFTFKFQFNPNRSLIK2ITQLLG-RTHATGIRKVRLEFN 224
D 206 PTKQVGLIQYANNPRVFNLTNTYKTBEMIVATSTQSVQGLINTTGAQVARYAYS 265
QY 225 ITNGARKNAPKILITDGEKFDGPDLYEDVPEADREGVIRVIVG-----GDAFRSEK 279
D 266 AASGGRSATKVVVVVVDGESH-DGSMILKAVIDOCNHDNILRFGIAVLGYLNALDITKN 324
QY 280 SROELNTVASKPRDHVFOINNFEALKTIONOLREKIFAIEGTQTGSSSSFFHEMSQEGF 339
D 325 LIKEIKAIASIPTEYFFNVSDAEALLKAGTIGEQIFSIETGVQ-GDNFQEMSQVGF 383
QY 340 SAAITNGP--LLSTVGSYDMAGEVFLYTSKEKSTFINMT--RVDSDM-DAYLGYAAAI 394
D 384 SADYSSONDILMAGVAGFAGWSGTIVKQTSKSHGLIFPKQAFDQILQDRNHSSVLYGSVAA 443
QY 395 ILNRVQSLVLAGPRYCHGLVAKFQNTGMWESNAV-----KGTQIGAYTGAISCSV 448
D 444 ISTGESTHVFAGAPRANYTQIIVLYSVN-----ENGNIIVIAHAGDQIGSYFGSLCSV 498
QY 449 DVDSNGSTDLVLIGAPHYEYQTR--GQOVSVCPLPRGORARWOCDAVLYGEQOPWGRFG 506
D 499 DVDKDTITDVLVAGAPWMSDLKKBEGRVYLFITKIGILGQH---FLEGPEGIENTRFG 555
QY 507 AALTVDGVNGDKLTVAIGAPCEENRGAVLFGTSGSGISPSHSORTAGS--KLSPR 564
D 556 SAIAAALSDINMGDFNDVIGSPLENQNGSVAIYINGHQT-IRTKYSQKILGSDGAFESH 614
QY 565 LQYFGSLSGGQDLTMDGLVLTGAGHYLLRSQPLRVKALMEFPNPREVARNVEFN 624
D 615 LQYFGRLDGYGLNGDSITDVSIGAFQGVVQLWSQSIADVAIEASTPEKI--TLVKN 672
QY 625 DQVVKGEAGEVRVCLVQKSTDRLEBGOIQSVVYTDALD---SGRPHSAVFNVEK 680
D 673 AQII-----LKLCP---SAXPRPKQNNQVAIVNITLDDAGFSRVTSRGLFKENN 721
QY 681 NSTRRTQVLGLTQC--ETLKLQPNCEIDPSPVIVRLNFSIANGTPLSAFGLNRLPVL 738
D 722 ERLCLQNMVNVQASCEPHIYIQES---DVNSLDLRVDISLENPGTS-----PALE 772
QY 739 EDAQRLFTALPPEKNGCNDNIQDDLSITF---SPMSLDCLVVGQFPRNVTVTVRND 794
D 773 AYSETAKVFSIPFHKDCGEDGLCISDLVDVRQIPAAQEQPFIVSNQNKELTFSTVLKN 832
QY 795 GEDSYRTQVTFPPLDLISYRKVSTLQNRQSRWLACESAST-EVSGALKSTSCSINH 853
D 833 RESAYNTGIIVDFSENLPF-----ASFSLPVDGTEVTCQVAAQSKSVACDVGY 880
QY 854 PIPENSEVTNITPFDVDSKASLGNKLLKANVTSENMPRTNKTEFQLELPVYATVMV 913
D 881 PALKREQQVTFITNFDNLQ-NLQNOASLSFQALSSEOBENKADNLVNLKIPLLYDAEI- 938
QY 914 VTHGVSITKVLNFTASENTSRVMOHQVQVSNLGR-----SLPISLVFIV---- 958
D 939 ---HLRTSTNINFEISSDGNVFSIVHSFEDVGPKTFSLKVTGSPVSVNATVILHIPO 995
QY 959 -----EVRNLQTVIWDPRPQVTF-SBNLSSSTCHTKER 988
D 996 YTKENPLMVLTVGTQTKAGDISCNADINPLKIGQT-----SSSVSPKSENFR---HTKE- 1047
QY 989 LPSHSDFLAELRKAPVNVCSIANQRIQCQDIPFGIQEBENATLKNLSFDWYIKTSHNH 1048
D 1048 -----LNCRTASCNVTCWLKDVHMKGEFVNVTTIRWNGTFASSTFQT 1091
QY 1049 LLIVSTAEI-LFNDVSFTLLPGQGAFFVRSQETKVEPFEPNP-----LP--LIVGSSVG 1100
D 1092 VQLTAAAEINTYNPPIVVI-----BDNTVTIPLMIMKPKDEKAEVPTGVIIGSIIA 1141

QY 1101 GLILLALITRAALYKLGFFKQYKDM 1125
 DB 1142 GILLILALVAILKLGFFKQYKDM 1166

RESULT 11
 A35854
 Integrin alpha-1 chain precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 23-Oct-1990 #sequence_revision 13-Sep-1991 #text_change 15-Sep-2003
 C:Accession: A35854; S11243
 R:Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto, J. Cell Biol. 111, 709-720, 1990
 A:Title: Molecular cloning of the rat integrin alpha-1 subunit: a receptor for laminin A
 A:Reference number: A35854; MUID:90338125; PMID:2380249
 A:Accession: A35854
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1180 <IGN>
 A:Cross-references: GB:X52140; NID:G56493; PIDN:CAA36384.1; PID:G56494
 C:Keywords: cell adhesion; cytoskeleton; transmembrane protein
 P:170-345/Domain: von Willebrand factor type A repeat homology <WMA2>

Query Match 18.1%; Score 1062; DB 2; Length 1180;
 Best Local Similarity 27.4%; Pred. No. 2.9e-65;
 Matches 344; Conservative 201; Mismatches 478; Indels 234; Gaps 47;

QY 1 FNLDTENAMTFOENARG-FCQSVVQL---CGSRVWVGAPQEIIVAAQNRGLYQCYSTGS 56
 DB 29 FNVDKNSMSPGVEDMFGYTVQYENBEGKVLGSPVVGQPKARTGDVYKCPVGR 88
 QY 57 CEP-IRLQVPVEA-----VNMSLGLSLAATSPQQLACGTPVHQCSENTRYKGL 106
 DB 89 AMPCKVLDLPVNTSIPNVTEIKENMTFGSTL-VTNPGGFLACGPLYAVRCGHLHYTTGI 147
 QY 107 CFLPGSNLRQOPKQPEALAGCQEDSDIAFLVDGSGSIIPHDFRAKKEPISTVMQLK- 165
 DB 148 CSDVSPTFQVNSNAP--VQECSTQ-LDIIVLDGSGNSIYP--WESVIAFLNLLKMDI 202
 QY 166 -KSTFLSFMQYSEPRIFHTFFEFQNNPNRSLIKPIQLLG-RHTATGIRKVVRELF 223
 DB 203 GPKQTQVGIQYGVNTHFENLANKYSTEEVLVAANKIGRQGLQWTALGIDTARKEAP 262
 QY 224 NITNGARKNAFKLILITDEKFGDPLGYREDVPEADRGVIRYVIGVDAPR-----SE 278
 DB 263 TEARGARGVKVMVITVDGESH-DNYRLKQVQDCDENIQRFSAIALGHYRGNLSTE 321
 QY 279 KSRQELMTVASKPRDHVFOINNPALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEG 338
 DB 322 KVEEIKSIASEPTEKFFNVSDALVTIVKALGERIFALEATADQSAASFENMSQTG 381
 QY 339 PSAAITNGPLLSVTGSDYDAGGVFLYTSKEKSTFINT--RVDSMDND---AVLGYAAA 393
 DB 382 FSAHYSQDVMVLMGAVGAYDNGTVVMQKQNMVIPHNTTFTQTEPAKNEPLASLYGTYN 441
 QY 394 IILNRRVQSLVLAGPRYOHIGLVAMPONTGMWESNANVGTQIGAYFGASLCSVDVDSN 453
 DB 442 SATIPGDVLYTAGOPRINHTQVVIYKXEDGNINILQTLGGEQIGSVFGSVLTIIDKD 501
 QY 454 GSTDLVLIGAPHY-----YQYR-GGQSVVCPIPRGORARWQCDVLYGE 497
 DB 502 SYTDLLLVAGPMYMGTEKEQGVVYVAVNQTRPEYQMSLEPIRQTCSSILKONSCTKEN 561
 QY 498 QGQPWG-RFGALVTLVGVNCKLTDVYAGPGEEDNRGAVYLPFGTSGSGISPSHSORI 556
 DB 562 KNEPCGARFGTAIAVXDLNDVGDNDVVIAGPLDDEHAGAVIYHSG-KGTIREATAORI 620
 QY 557 AGSKLSRLQVFGQSLGGQDLTWDGLVLTVTGAQGHVLLRSQPLVRVKAIKFEPREV 616
 DB 621 PSGGDKTLKFFGQSIHGEDMNGDGLTDTVIGLGGAAFWARDVAVVAVVWNNFENKV 680
 QY 617 ARNVFECNDQVVKKEAG--EVRVCLHVO-KSTFDRLREGQISVWVYDLALDSGRPHSR 673

DB 681 NIQKNCR---VEGKETVCINATWCFFHVKLAKEDSIYEADLQ-----YRVTLDSLRLQISR 733
 QY 674 AVFNET-----KNSSTRQTOVLGLTQTCETLKLQLPNCI-----EPVSPVIVLRL 718
 DB 734 SFFSGTQERKIQRIITVRES-----CIRHSFVMLDKHDFQDSVRVTL 776
 QY 719 NFSLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPFFKNCNGDNICODDLSTITFSFMSLCLIV 778
 DB 777 DNLTL-DENG-----FVLDLALPNSVHEHIPFAKDCGNKERCISDLTLNLTSTTEKSLLI 830
 QY 779 VGGPRE-FNVTVTVNDGEDSYRTQVTFPFFLDLSYRKVSTLQNRORSORWELACESASS 837
 DB 831 VKSQHDKFNVSILTVKNGDSAYNTRTVVQHSFNLIFSGBIEIQKD-----SCSEN-- 880
 QY 838 TEVSGALKSTSCSINHPIFPENSEVTENITPDVDSKASLGNKLL-LKANVTSENNMPTN 896
 DB 881 -----QNTICRVGYFPFRAGETVTFKLIIFQFVTSLSSENAIHLSATDSEPLESLN 933
 QY 897 KTEFOLELPKYAV---YMWVTSHGVT-----KYLNTFASNTSRVMOHQVQVSNL 945
 DB 934 DNEVNISIPVKYEVGLQFYSSASEHHISVAANETIPBFINST--EDIGNENINVFYTIKRR 991
 QY 946 GORSIP---ISLVP-----LVPVRLNQTVIWD-----RP-----Q 972
 DB 992 GHFPMPELQLSISFPNLTFADGYPVLYPIG-----WSSSDNVNCRPSLEDPFGINSKK 1045
 QY 973 VTFS-----ENLSSTCHTKERLFSHSDFLAELRKAPVNVCSIAVCORIQCDI--- 1019
 DB 1046 MTISKSEVLKRTGIQDCSSTC-----GVATITCSLLPSDLQSVNVSILL 1088
 QY 1020 ---PFFGQEEP---NATLKGNLSDWYIKTSHHLLIVSTAELFNDSVFTLLPQGNF 1073
 DB 1089 LMKPTF-IRAHFSSLNLTLRGELK-----SENSLTSLSN----- 1123
 QY 1074 VRSQTEKVEPEVPNPPLP---IVGSSVGGLLLLALITAAALYKLGFFKQYKDMSE 1128
 DB 1124 RXBELAIQISKDGLPQRPVPLWILLASAPAGLLMLLILALWKIGFFKPLKKKMEK 1180

RESULT 12
 A41131
 Lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse
 N:Alternate names: Integrin alpha-4
 C:Species: Mus musculus (house mouse)
 C:Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 24-Sep-1999
 R:Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzmann, B.; Weissman, I.L.
 J. Cell Biol. 115, 1149-1158, 1991
 A:Title: Cloning and expression of cDNAs for the alpha subunit of the murine lymphocyte
 A:Reference number: A41131; MUID:92064645; PMID:1840602
 A:Accession: A41131
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1039 <NEU>
 C:Cross-references: EXBL:X53176; NID:G51484; PIDN:CAA37316.1; PID:G51485
 C:Superfamily: Integrin alpha-4 chain
 C:Keywords: cytoskeleton; transmembrane protein

Query Match 11.3%; Score 663; DB 2; Length 1039;
 Best Local Similarity 22.7%; Pred. No. 1.5e-37;
 Matches 281; Conservative 187; Mismatches 403; Indels 368; Gaps 48;

QY 1 FNLDTENAMTFO-ENARGFGQSVV-QLQGR--VVVGAPQEIIVAAAN---QRGSLYQCYD 52
 DB 41 YMLDPENALLYQGPSOTLFGYSVWLHSHGSKWLVGAPTASWLSNASVNVNPGAIYRCGI 100
 QY 53 STG---SCPIRLQVP-----VEAVMSLGLSLAATSP-PQLAACG---PTVHQ 95
 DB 101 RKNPNQTCQLQSGSPGSGPCGKTCLIEDNOMLVTLRSQPGNGSVITCGHRWKNIFY 160
 QY 96 TCSNTYVAGLCLFGSNLRQOPKQFPPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKE 155
 DB 161 MRSNDKLPVGIYCVWPSDLRTLSK----- 185

```

156 F1STVMEQLKKSXTLFSLMQYSEEFRIHFTTKEFQNNPNRSLKPIQTOLLGRTHRTATGI 215
186 -----RMAPCYKDYT-----
216 RKVURELFNITWGARKNAFKLILITDKEFGDPLGYEDVIEADREGVIRVIGVGDAF 275
196 -----RKGE-----
276 RSEKSRQBELNTVASKPPDRHVFQIINNFRALKTIQIOLREKIFAIEGTQTSSSSFEHMS 335
201 -----NFAS-----C 205
336 QBEGSAATISNGPLLSITVGSYDWAGGVLY--TSKEKSTFINMTRVDSMDANDAVILGYA-- 391
206 QAGISSTPYTODLIVMGAPGSSYWTCTVFNYNITNQYKAFVD--RQNVQKTSGLYGSVG 263
392 AAIIILNRVQSILVIGAPRYQHIHGLVAMPQONTGMMESNANV--KZQIQAIGYFASLCS 447
264 AGHFRSPHTTEVVGAPQHEQIGKAYIP---SIDENELNIYEMKKGKLSYFYGASVCA 319
448 VDVDSNGSDTLVLIAGAPHYEEQTRGGQSVCLPRGQRA--RWCQDAVLYGEOQPWGRFG 506
320 VDLNADGFSDL-LVGAPMOSTIREEGRFVY--INSGWGVYMERVLVGSBKYA--ARFG 376
507 AALTVLGVNGVKLDTVAIGAPEEDNRGAIVLPHGTSGSGLISPHSORSIAGSKLSPLIQ 566
377 ESIANLGIDINDGFEIAIQAQEDDLRGAVIYNGRV--DGSISSYSORIEQQIKSLSR 435
567 YFQSLSGGQDLTMOGLVDLTVGA--QGHVLLRSQPLVRKAIMEFNPREVARNVFECN 624
436 MFGQSIISGQIDADNNGVVDVAGAFQSDSAVLLRTPFVIVEASLS--HPESVNRKFPDCT 494
625 DQWVKKEAGEVAVCLHVQKSTRDLREQIQSVTYDIALDSGR---PHSAVP--NET 679
495 -----ENGLPSVCWHLTLCPYSKQKVPGYIVLFPYVSLDVHRKABSPSRPYFSSNGT 547
680 KNSTRQTOVLGTQTCETLKLQLPNCIDBPSPVILRLNFSI-----VGTPLSAFGNLR 734
548 SDVITGISRVSSSGEKCRTHQAFMRQVDRLTPIHVEATYHLGHVITKRNTESPPPLQ 607
735 PVLAEDMOR-LFTALPPEKNCNDNI CODDLSITFS-----FMSLDCLVVGGREFNV 787
608 PILQOKKBDVIRMTNFAFCAYEN-CSADLQVSAKVGLFKPYENKTYLAVGSMKTHML 666
788 TVTVRNDGDSYRTQVTFPPPLDLSYRKYSTLQNRQSQRWLACESASSTEVSGALKST 847
667 NVSLFNAGDDAYETLLNVQLPTGLYFIKILDSEK-----QINCE---VTSESSGIVK-L 716
848 SC SINHIPIPNSEBVTNITFDVDSKASLGNKLLKANVTSEN--NMPRNTKTEFOLELP 905
717 ACSIGYTYVDRLSRIDISFILLDVSSUSRAHEDUSIVHASENEGLEDQVRDNRTVLTIP 776
906 VKAVYVMVATSHGV--STKYNLNFASNTSRVMHQ-----YQVSNLGRSLP-ISLVFL 957
777 LRREV--MLTVHGLVWFTSVYSGSSENEPETCMAEKILNTPHINTGISMAPNVSVKIM 834
958 VPVELNQTVIWD RP--QVTFSENLSSTCHTKE-----RLPSHSDFLAEALR 1000
835 VP---NSFLPODDKFLNVLDVQTITGQCHPKYHGRECTPAQQKIGATLTDIVKFLSKTD 891
1001 KAPVVMGSLA--VCQRIQCDDIPPEGIOEFPNATLKGNSLFDWTKTSHNHLIIVSTAEL 1058
892 KR-LLYCMKADQHCLDFLCN--FGWMESG-----KEASVHIQLEGRPSIL 933
1059 FNDVSFTLLPQOGAFVRSQSTETKVEPEVNP----- 1090
934 EMDETSIL-----XFEIKATAPPEHPKPVIELKNDENVAVTSLGHLHQRPKHF 983
1091 -LPLIVSSVGGLLILALITAALYKLGFFKFRQYKDWSE 1128
984 TIIITISFAGLIVHVAISCMWKAGFFKFRQYKSILOE 1022

```

RESULT 13

S06046

integrin alpha-4 chain precursor - human

N;Alternate names: CD49d; very late antigen-4 alpha chain; VLA-4 alpha chain

C;Species: Homo sapiens (man)

C;Date: 01-Dec-1989 #sequence revision 01-Dec-1989 #text_change 24-Sep-1999

C;Accession: S06046; A39355; D28018

R;Takada, Y.; Elices, M.J.; Crouse, C.; Hemler, M.E.

EMBO J. 8, 1361-1368, 1989

A;Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other integrin

A;Reference number: S06046; MUID:89356603; PMID:2788572

A;Accession: S06046

A;Molecule type: mRNA

A;Residues: 1-1038 <TA>

A;Cross-references: GB:X16983; EMBL:X15356; NID:G33945; PIDN:CAA34852.1; PID:G33946

R;Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.

Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991

A;Title: Characterization of the alpha-4 integrin gene promoter.

A;Reference number: A39355; MUID:91239513; PMID:2034655

A;Accession: A39355

A;Molecule type: DNA

A;Residues: 1-93 <ROS>

A;Cross-references: GB:M62841

R;Takada, Y.; Strominger, J.L.; Hemler, M.E.

Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987

A;Title: The very late antigen family of heterodimers is part of a superfamily of molecules

A;Reference number: A94151; MUID:87204112; PMID:3033641

A;Accession: D28018

A;Molecule type: protein

A;Residues: 40-50, 'E', 52-53 <TA2>

C;Genetics:

A;Gene: GDB:ITGA4; CD49D

A;Cross-references: GDB:I128032; ONIM:I192975

A;Map position: 2q31-2q32

C;Superfamily: integrin alpha-4 chain

C;Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; transmembrane

E;1-39/Domain: signal sequence #status predicted <SIG>

E;40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>

Query Match 10.8%; Score 635; DB 2; Length 1038;

Best Local Similarity 25.3%; Pred. No. 1.3e-35;

Matches 249; Conservative 154; Mismatches 370; Indels 212; Gaps 39;

QY 272 GDAPRSEKSRQELNTVASKPRD-----HFVQINNFEALKT-----IQN 310

DB 121 GKTCLERDQNLGVTLSRQFGENGSIIVTCGRHWKNIIFYIKNENKLPGTGCYGVPPDLRT 180

QY 311 QLRKI-----FALEGTTGSSSFEHMSQEGPSAAITNSGPIILSTVGSY 356

DB 181 ELSKRIAPCYQVYVKKGGENFA-----SCQAGISSFTTKDLIIVGAPGSS 225

QY 357 DWAGGVFLY---TSKEKSTFINMTRVDSMDNDAYLGVA--AAIILNRVOSLVLAGAPRY 411

DB 226 YWTGSLFVYNIITNKYKAFDKQNVKF---GSVLGYSVCGAGHFRSQTHTTEVVGGAPOHE 282

QY 412 HIGLVAMPFRONTGWESENANV---KQTQIAYFGASLCSDVDSDNGSTDLVLIGAPHYY 467

DB 283 QIGKAVIYF---SIDELKELNILEMKKGLGSYFGASVCAVDLNAUGFSOL-LVGAPMQS 337

QY 468 EQTRGGGVSVCPILPGRQARWQC-DAVLYGEQOGMGRFGAALTVLGDVNGDKLITDVAIG 526

DB 338 TIREGRVFVY-INSGSAVNMAMETNLVGSDDKA-AKFGEIIVNLGDIIDNDGFEDVAIG 395

QY 527 APGBEDNRGAVYLFHGTSGGSIIPSHSQRIAGSKLSRLOYPFGQSLGGQDLTMDGLAVDL 586

DB 396 AFQEDDLQGAIIYINGRA-DGISSTPFSQRIEGLQISKLSNFGQISGQIDADNNGVYDV 454

QY 587 TVGA--QGHVILLRSQPLVRKATLMEFNPREVARNVEPCNDQVVKGEAGEVRVCLHVQK 644

DB 455 AVGAFRSDASVALLTRPVPWIVDASLS-HPESVNRKEDC-----VENGMFSCIDLTL 506

QY 645 STRDLREGOIQSVVYTDYDALDSGR-----PHSRVAVNETKNSTRQTVLGLTQTCETL 699

```

Db 507 CFSYKGVGVYVLPYNSLDVNRKAEPPRPFPSSNGTSDVITGSIQVSSREANCRT 566
QY 700 KIQLPNCIEDPVSPIVLRNFSLVGTPLS-----AFGNLRPVLAADQR-LFTALFPPEK 753
Db 567 QAFMRXDVRLDILPIQIEAAYLGHPIVSKRSTEBFPPLQILOQKXDKIMKKTINFAR 626
QY 754 NCQNDNIQDDLSIT--FSPM-----SLDCLVVGGRPEFNVTVVRNDGSDSYRTQVTFPP 807
Db 627 FCAHEN-CAADLQVSAKIGFLKPHENKTYLAVGSMKTLMLNLSLNFAGDDAYETTLHVKL 685
QY 808 PLDSVRKVTSTIQNQRSORSLACESASSTEVSGALKSTSCSINHPIPPENSEVTFNIT 867
Db 686 PVGLYFIKILEEEK-----QINCE---VTDNSGVVQ-LDCSIGIYYVDHLSDIDISFL 735
QY 868 FVDSKASLGNKLLKANVTSEN--NMPRTNTEFOLELPVKYAVVMVTSHGVTGKYLIN 925
Db 736 LDVSSLRAEEDLSIVHATCENEEMDNKHSRTVAIPLKYEVKLTVHGFWNTSFY 795
QY 926 FTASENTRSV-----MHOYQVSNLQGRSLP-ISLVFLPVVRLNQTIVWDRPQVTFSEN 979
Db 796 GSDNEPETCMVEKKNLTFHVINTGSMAPNVSVBEIMVENSFS-----PQTDKLFNI 848
QY 980 -----SSTCHTKERLPSSHDFLAEIRKAPVNVCSIAVCQRIQCDIPRFGIOBEFNATLK 1033
Db 849 LDVQTTTGCHFENYQ-----RVCALEQ-----QKSAQOTLK 880
QY 1034 GMLSFOWYIKTNSHLLIIVSTAB-----ILFN-----DSVFTLLPGQAFVRSQFE 1079
Db 881 GIVRF--LSKTDKRLLYCIKADPHCLNFCNFGKMSGKESVHIQLEGPS-ILEMDE 936
QY 1080 TKVEPFEV-----RNP-----LPLIVGSSVGGLL 1103
Db 937 TSALKEFEIRATGPEPNPRVIELNDENVAHVLEGLHQRPKRYPTIIVSSLLGLI 996
QY 1104 LLALITAAALYKLGFFRKQYKDMSE 1128
Db 997 VLLLIISYVMWKAGFFRKQYKSIILQE 1021

RESULT 14
I58409
Integrin alpha-9 chain precursor - human
C:Species: Homo sapiens (man)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 31-Dec-2000
C:Accession: I58409; A49459
R:Hibi, K.; Yamakawa, K.; Ueda, R.; Horio, Y.; Murata, Y.; Tamari, M.; Uchida, K.; Takah
Oncogene 9, 611-619, 1994
A:Title: Aberrant upregulation of a novel integrin alpha subunit gene at 3p21.3 in small
A:Reference number: I58409; MUID:94119603; PMID:8290272
A:Accession: I58409
A:Status: preliminary; translated from GB/EMBL/DDSB
A:Molecule type: mRNA
A:Residues: 1-1035 <RES>
R:Cross-references: GB:D25303; NID:9464180; PIDN:BA04984.1; PID:9533327
R:Palmer, E.L.; Ruegg, C.; Ferrando, R.; Pytel, R.; Sheppard, D.
J. Cell Biol. 123, 1289-1297, 1993
A:Title: Sequence and tissue distribution of the integrin alpha9 subunit, a novel partne
A:Reference number: A49459; MUID:94064789; PMID:8245132
A:Accession: A49459
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 30-1035 <PAL>
R:Cross-references: GB:I24158
C:Superfamily: integrin alpha-4 chain
C:Keywords: glycoprotein; metal binding; transmembrane protein
P:1-27/Domain: signal sequence #status predicted <SIG>

Query Match 10.7%; Score 630; DB 2; Length 1035;
Best Local Similarity 26.5%; Pred. No. 2.9e-35;
Matches 237; Conservative 160; Mismatches 332; Indels 164; Gaps 40;
QY 331 EHEMSQGFSAITSNGPLLSITGVSYDWAGVGLYTSKESKSTFINWT-RVDSMDNDAYLG 389

```

```

Db 190 EHGSCQAGIAGFTTEELVWVGAPGSPYAGTGIKVLNLT-D-NTYKLKLNDEVIMRRYTYLG 248
QY 390 YA--AAIILNRVQSLVGLGAPRYOHIGLWAFR--QNTGMWESNANVKGTOIGAVFGASL 445
Db 249 YAVTAGHPSHSTTDVVGAPQDGIKGVIFRADRSGLIKIFOASGKKMGSYFGSSL 308
QY 446 CSVDVDSNGSTDLIGAPHYYEQTGGQVSVCPPLFGQARMOCDVAVLYGEOQGPWGRF 505
Db 309 CAVDLNGDGLSDL-LVAGAPMFSEIRDEQVTVY--INRGNGALEE-QLALTGD-GAYNAHF 364
QY 506 GAALTVLVDYNGDKLTVAIGAPGEEDNRGAVLPHGTSGSGISPSHSORISAGSKLSPL 565
Db 365 GESLADLDDNDDGFPDVAIGAPKEDDPAGVAYIYHGDA-GIYVQYMKLSGQKINPVL 423
QY 566 QYFQSSLSGGDLTMDGLVDLTGCA--QGHVLLRSQPLVRKAKMFPNPRVARNPFC 623
Db 424 RMFGSISGGIDMDGNGYEDVTGAFMSDSVLLRARPVITVD-VSIFLPGSINITAPQC 482
QY 624 NDQVVKGEAGEVAV-CLHVQ-----KSTRDLRBEQIQSVVYVYDLDLSDGRPHSAV 675
Db 483 HD-----GQQPVNCLNVTTCSFHGKHVPBEI---GLNYVLMDADVAKKEKQOMPVY 531
QY 676 F---NETKNSTRRTQVIGLGTQCTETLKQLPNCIEDPVSPIVLRNFSL-----VGTPLS 728
Db 532 FVLLGETMGQVTEKLQTYMEETCRHYVAHVRRVQDVISPIVFEAAVSLSEHVGBEER 591
QY 729 AFGNLRPVL-----ABDAQRLLTALFPPEKNGNDNTICQDLSITFSPM--SLD-----CL 777
Db 592 ELPLTTPVLRWKKGQKIAQKQTV---FERNCRSED-CAADLQGLQKLLSSMDKTYL 647
QY 778 VVGPRBNPVTVVRNDGEDSYRTQVTFPFDLSYRKVSTLQQRORSORSLACESASS 837
Db 648 ALGAVKNISINISINLGDADYDANVSVNSRELFF-----INWQKEENGISCLEL 701
QY 838 TEVSGALKSTSCSINHPIPPENSEVNTITPDVDSKASLGNKLLKANVTSENMPRINK 897
Db 702 DPL-----KCSVGFPPMRSKSYFSPVDPDTHLS--GEEVLSPVITVTAQSG--NTER 750
QY 898 TE-----FQLELPVKYAVVMVTS-----HGVSTKYLNFASENTS---RVMQHY 940
Db 751 SESLHDNTVLNVLMPLMHEVDTSITGIMSPTSFVVGESVDAAFIQLDDLECHQFQINITL 810
QY 941 QVSNLQGRSLPISLVL-VPVRLN-----QTVIMDRPQVTPSENLSSTCHTKER 988
Db 811 QVNTGPTPLGSSVSISFPNRLSSGGAEMFHVQEMVVGQEKNGCSFOKNFTPCIIPOEQ 870
QY 989 LPSSDPLAEIRKA-----PVNCSIAVCQRIQCDIPFGIOBEFNATLKGNLSFD 1039
Db 871 ENIFHTIIFATFKSGRKVLDCCKPGISCLTAHCN-----PSALAKESRTI-----D 917
QY 1040 WYIKTSHNHLIIVSTAEILFNDSVFTLLPGQAFVRSQTEKVEP----- 1084
Db 918 IY-----MLNT-ELIKXSSSVIQ-----FMSRAKVQVDPALRAVVEIAHGNPEEV 962
QY 1085 ---FEVFN---PLPLIVG-----SSVGLLILLALITAAALYKLGFEKRYKDM 1126
Db 963 TVVPEALNLEPRGVYVGVWIIAISLLVGLIFLLAVLLWKGPFRRIRKII 1015

RESULT 15
T31437
Integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)
C:Species: Lytechinus variegatus (variegated urchin)
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C:Accession: T31437
R:Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A:Description: Alpha SU2, a sea urchin integrin which binds laminin.
A:Reference number: Z21035
A:Accession: T31437
A:Status: preliminary; translated from GB/EMBL/DDSB
A:Molecule type: mRNA

```

A:Residues: 1-1041 <HER>
A:Cross-references: EMBL:AF067658; MID:g32202040; PID:g32202041; PIDN:AAC23572.1
A:Experimental source: developmental stage embryo
C:Function:
A:Description: binds laminin
C:Superfamily: integrin alpha-2b chain

Query Match 10.5%; Score 614.5; DB 2; Length 1041;
Best Local Similarity 24.8%; Pred. No. 3.5e-34;
Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

QY 336 QEGFSAATSNGP--LLSTVGVSDWAGVFLYTSKEKSTFINMTRVDS-----DMNDAYLG 389
DB 181 QAGFGIIFSDNSALVMGAPGYLGGYIYQSLNRSV-VQTOESNTGYTSPDNSTYRG 239
QY 390 YAAAI--ILNRVQSLVIGAPRYQHI-GLVAFRONTGMWESNANYKGTQIGAYFGASIC 446
DB 240 YSLALGDFNGDGVDYVVGTPRAESLGLVAIFDQNLQFN--QVMGTQIVAYFGYSVT 296
QY 447 SYDVDSNGSTDLVIGAPHYGTGCGQSVSCPLPGQARWQCDVLYGEQ----- 498
DB 297 VVDI-NNDTYDILLVGAPEYMDGPAIQ-----RWEAGAVYYVLQNPDPVGCA 343
QY 499 -----GQWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRCGAVYLFHGTSGS 546
DB 344 SNRLSSTLIGQIRSRFGLSIASIGDSNQGDFNDVAIGAPYEGDDAGAVIYHG-SAN 402
QY 547 GSPSHSQRIAGSKLS-PRLOVFGSLSGQDLTMDGLVDLTVGAQ--GHVLLRSQPVL 603
DB 403 GLKSTPAQVLTPTSLGSHGITTFPSLQGGQMDKNKYPDLLWGAESANTAVLIRTPVY 462
QY 604 RVKAIMFEFPRVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLRGQIQSV----- 658
DB 463 SLDATLNTPE-----IGINLENKTYE-LADGTWVTSFIAMT 497
QY 659 -----VYVDLALDSG-RPHSRVAFNETKNSRQTVQVGL-TQTCETLKLQ 702
DB 498 CFTYTGNYLPDHIDISVTVVDSIIANRRAMPVDMDSBITKTRRLAVSTQFCDEPLAY 557
QY 703 LPNCIEDPSPVLRNPLSVLGTPLSAPGN-----LRPVLAEQAQLPTALPPEK 753
DB 558 VGNSEDKLPPIKVTQYDL-----NDESRLOPHEILPDIIDNATSTQTKQVSION 609
QY 754 NCGNDNIQQDLSITTFMSLDCLVGGPREFNVTVVRDGEDSYRTQVTFPFLDLSY 813
DB 610 NCYN-NICIFDLDVTVT-PNLPNVIGQTELTLDSVLSNRRGEDAFQSSLSVYVPLQOF 667
QY 814 RYVSTLQNRQSRWRLACESASSTEVSGALKSTCSINHPIPEX-----SEVTFNIT 867
DB 668 VRL-----ERKANMDPSVTCSEDSO-----LRIITCDTGNPMVGNKILFGLTLSTFQVS 717
QY 868 FDVDS-----KASLGNKLLKANVTSNNSMPRTNKTEFQLELPVKYAVYVMTSHGVSTK 922
DB 718 GDKDSIBFYFKASENS--EDPNTLENLNMVTVTVDCITLKLSSYPEIVMYSYTOED 775
QY 923 YL-----NFTASENTSRVWQHOYQV-----SNLQGRSLPIS-----LVFLVPV 960
DB 776 YVVPVPPAKNASADIGMEVW-HLYEVNRNTGSSNAGEVSLNIQWPKNEDGEYLFYLLGI 834
QY 961 RLNQTVLWDRPO-----VTESENLSSTCHTKERLPSHSDFLAELRKAPVY 1005
DB 835 MTEEGVTCQLTQKANEPEGVLAPSTYAKLSNSTQVSGRKRREPEVAEALQTDN--VI 892
QY 1006 NCSTAVCQRIQCDDIPFGIOEEFNAT-----LKNLSFDWYIKTSHNHLIVSTABILF 1059
DB 893 YCASDSCVLINCTI-----DEINASKSKVVRILGRF--W-----ERTF 928
QY 1060 NDSVFTLLPGGAPVRSQETKYE--PFEVDPNP-----LP----- 1092
DB 929 QKAVSELTVPVQATIASASAANKTIPYNIPLRDFSDSTKASTLVTEELVPPVPIAW 988
QY 1093 -LIVGSSVGLLILALITALYKLGPFKR-----QYKDMMS--EGGPP 1132

Db 989 WIIVSVLGGIILLIILGLWKCGPFRKKPGKEKEYAPVASADKGGPP 1038
Search completed: June 7, 2004, 17:18:08
Job time : 20.559 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:55:38 ; Search time 9.97194 Seconds
(without alignments)
5937.039 Million cell updates/sec

Title: US-09-902-481b-5
Perfect score: 5876
Sequence: 1 FNLDTENAMTQENARGFGQ.....FKRQYKDMWSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5846.5	99.5	1152	1 ITAM HUMAN	P11215 homo sapien
2	4470	76.1	1153	1 ITAM MOUSE	P05555 mus musculus
3	3459	59.0	1163	1 ITAX HUMAN	P20702 homo sapien
4	3411	58.0	1162	1 ITAD HUMAN	Q13349 homo sapien
5	1557.5	26.5	1170	1 ITAL HUMAN	P20701 homo sapien
6	1538.5	26.2	1163	1 ITAL MOUSE	P24063 mus musculus
7	1161.5	19.8	1167	1 ITAE MOUSE	Q60677 mus musculus
8	1153	19.6	1179	1 ITAE HUMAN	P38570 homo sapien
9	1103.5	18.8	1151	1 ITAL HUMAN	P56199 homo sapien
10	1093.5	18.6	1189	1 ITAH HUMAN	Q9ulx5 homo sapien
11	1085	18.5	1170	1 ITAZ BOVIN	P53710 bos taurus
12	1071	18.2	1178	1 ITA2 MOUSE	Q62469 mus musculus
13	1068	18.2	1181	1 ITA2 HUMAN	P17301 homo sapien
14	1064.5	18.1	1167	1 ITAG HUMAN	O75578 homo sapien
15	1062	18.1	1180	1 ITAL RAT	P18614 rattus norv
16	663	11.3	1039	1 ITA4 MOUSE	Q00651 mus musculus
17	635	10.8	1038	1 ITA4 HUMAN	P13612 mus sapien
18	630	10.7	1035	1 ITA9 HUMAN	Q13797 homo sapien
19	593.5	10.1	1032	1 ITA9 XENLA	Q91687 xenopus lae
20	571.5	9.7	1066	1 ITA3 CRISP	P17852 cricetidae
21	567.5	9.7	1053	1 ITA3 MOUSE	Q62470 mus musculus
22	555.5	9.5	1053	1 ITA5 MOUSE	P11688 mus musculus
23	545.5	9.3	1050	1 ITA5 XENLA	Q06274 xenopus lae
24	542.5	9.2	1034	1 ITAV CHICK	P26008 gallus gall
25	537.5	9.1	1130	1 ITA6 HUMAN	P23229 homo sapien
26	535	9.1	1044	1 ITAV MOUSE	P43406 mus sapien
27	532.5	9.1	1066	1 ITA5 HUMAN	P26006 homo sapien
28	532	9.1	1049	1 ITA5 HUMAN	P08648 homo sapien
29	531.5	9.0	1072	1 ITA6 CHICK	P26007 gallus gall
30	526	9.0	1048	1 ITAV HUMAN	P06756 homo sapien
31	517	8.8	1044	1 ITA8 CHICK	P26009 gallus gall
32	512.5	8.7	1091	1 ITA6 MOUSE	P61739 mus musculus
33	498	8.5	1179	1 ITA7 MOUSE	Q61738 mus musculus

ALIGNMENTS

RESULT 1

ID	ITAM HUMAN	STANDARD	PRT	1152 AA.
AC	P11215			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) {Leukocyte adhesion receptor M01}			
DE	(Neutrophil adherence receptor).			
GN	ITGAM OR CR3A OR CD11B.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein M01: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor M01 (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.P., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098893; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).			
RN	[6]			

34	496	8.4	1396	1	ITA2 DROME	P12080 drosophila
35	491.5	8.4	1146	1	ITAL DROME	Q24247 drosophila
36	491	8.4	1033	1	ITAB_MOUSE	Q8qum0 mus musculus
37	489.5	8.3	1039	1	ITAB_HUMAN	P08514 homo sapien
38	489	8.3	1025	1	ITAB_HUMAN	P53708 homo sapien
39	486	8.3	126	1	ITAM_CAYPO	P15578 homo sapien
40	476	8.1	1181	1	ITA7_HUMAN	Q13683 homo sapien
41	470	8.0	1106	1	ITA7_RAT	Q63258 rattus norv
42	467	7.9	1226	1	PAT2_CAEEL	P34446 caenorhabdi
43	445.5	7.6	1139	1	INAL_CAEEL	Q03600 caenorhabdi
44	424	7.2	1115	1	ITAS_DROME	O44386 drosophila
45	385	6.6	1000	1	ITAS_DROME	Q9wlm8 drosophila

RP	SEQUENCE OF 1-9 FROM N.A.	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
RP	MEDLINE=92073318; PubMed=1683702;	CC	the European Bioinformatics Institute. There are no restrictions on its
RA	Shelley C.S., Arnaut M.A.;	CC	use by non-profit institutions as long as its content is in no way
RT	"The promoter of the CD11b gene directs myeloid-specific and	CC	modified and this statement is not removed. Usage by and for commercial
RT	developmentally regulated expression."	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
RL	Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).	CC	or send an email to license@isb-sib.ch).
RN	[7]	CC	-----
RP	SEQUENCE OF 1-9 FROM N.A.	DR	EMBL; J03925; AAA59544.1; -
RP	TISSUE=Blood;	DR	EMBL; M18044; AAA59491.1; -
RP	MEDLINE=92144986; PubMed=1346576;	DR	EMBL; J04145; AAA59903.1; -
RA	Pahl H.L., Rosmarin A.G., Tenen D.G.;	DR	EMBL; S52227; AAB24821.1; -
RT	"Characterization of the myeloid-specific CD11b promoter."	DR	EMBL; S52152; AAB24821.1; JOINED.
RL	Blood 79:865-870(1992).	DR	EMBL; S52153; AAB24821.1; JOINED.
RN	[8]	DR	EMBL; S52154; AAB24821.1; JOINED.
RP	SEQUENCE OF 17-31.	DR	EMBL; S52155; AAB24821.1; JOINED.
RP	MEDLINE=87076671; PubMed=3539202;	DR	EMBL; S52157; AAB24821.1; JOINED.
RA	Pierce M.W., Remold-O'Donnell E., Todd R.P. III, Arnaut M.A.;	DR	EMBL; S52159; AAB24821.1; JOINED.
RT	"N-terminal sequence of human leukocyte glycoprotein Mol:	DR	EMBL; S52161; AAB24821.1; JOINED.
RT	conservation across species and homology to platelet IIb/IIIa."	DR	EMBL; S52164; AAB24821.1; JOINED.
RL	Biochim. Biophys. Acta 874:368-371(1986).	DR	EMBL; S52165; AAB24821.1; JOINED.
RN	[9]	DR	EMBL; S52167; AAB24821.1; JOINED.
RP	X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.	DR	EMBL; S52169; AAB24821.1; JOINED.
RP	MEDLINE=95171458; PubMed=7867070;	DR	EMBL; S52170; AAB24821.1; JOINED.
RA	Lee J.O., Rieu P., Arnaut M.A., Liddington R.;	DR	EMBL; S52173; AAB24821.1; JOINED.
RT	"Crystal structure of the A domain from the alpha subunit of integrin	DR	EMBL; S52174; AAB24821.1; JOINED.
RT	CR3 (CD11b/CD18)."	DR	EMBL; S52180; AAB24821.1; JOINED.
RL	Cell 80:631-638(1995).	DR	EMBL; S52181; AAB24821.1; JOINED.
RN	[10]	DR	EMBL; S52184; AAB24821.1; JOINED.
RP	X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.	DR	EMBL; S52189; AAB24821.1; JOINED.
RP	MEDLINE=96363671; PubMed=8747460;	DR	EMBL; S52191; AAB24821.1; JOINED.
RA	Lee J.O., Bankston L.A., Arnaut M.A., Liddington R.C.;	DR	EMBL; S52192; AAB24821.1; JOINED.
RT	"Two conformational states of the integrin A-domain (I-domain): a pathway for	DR	EMBL; S52203; AAB24821.1; JOINED.
RT	activation?"	DR	EMBL; S52212; AAB24821.1; JOINED.
RL	Structure 3:1333-1340(1995).	DR	EMBL; S52213; AAB24821.1; JOINED.
RN	[11]	DR	EMBL; S52216; AAB24821.1; JOINED.
RP	X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.	DR	EMBL; S52219; AAB24821.1; JOINED.
RP	MEDLINE=98362595; PubMed=9687375;	DR	EMBL; S52220; AAB24821.1; JOINED.
RA	Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.;	DR	EMBL; S52221; AAB24821.1; JOINED.
RA	Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrikson R.L.;	DR	EMBL; S52222; AAB24821.1; JOINED.
RA	Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.;	DR	EMBL; S52226; AAB24821.1; JOINED.
RA	Mutcher V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;	DR	EMBL; W76724; AAA58410.1; -
RT	"Cation binding to the integrin CD11b I domain and activation model	DR	PIR; A31108; RMHUIB.
RT	assessment."	DR	PIR; A31108; RMHUIB.
RL	Structure 6:923-935(1998).	DR	PDB; 1A8X; 17-JUN-98.
RN	[12]	DR	PDB; 1BHO; 18-NOV-98.
RP	3D-STRUCTURE MODELING OF 17-616.	DR	PDB; 1BHQ; 18-NOV-98.
RP	MEDLINE=98226734; PubMed=9560195;	DR	PDB; 1IDN; 25-NOV-98.
RA	Oxvig C., Springer T.A.;	DR	PDB; 1IDO; 01-AUG-96.
RT	"Experimental support for a beta-propeller domain in integrin alpha-	DR	PDB; 1JLM; 11-JAN-97.
RT	subunits and a calcium binding site on its lower surface."	DR	PDB; 1MIU; 07-AUG-02.
RL	Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).	DR	Genew; HGRC:6149; ITCAM.
CC	-1- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS	DR	MIM; 120980; -
CC	ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES	DR	GO; GO:0008305; C:integrin complex; TAS.
CC	AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.	DR	GO; GO:0004895; P:cell adhesion receptor activity; TAS.
CC	IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF	DR	GO; GO:0007155; P:cell adhesion; TAS.
CC	THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D	DR	InterPro; IPR000413; Integrin_alpha.
CC	PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR	DR	InterPro; IPR002035; VWF_A.
CC	FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES	DR	Pfam; PF01839; FG-GAP; 3.
CC	OF FIBRINOGEN GAMMA CHAIN.	DR	Pfam; PF00357; Integrin_A; 1.
CC	-1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M	DR	Pfam; PF00092; vwa; 1.
CC	ASSOCIATES WITH BETA-2.	DR	PRINTS; PR01185; INTEGRINA.
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.	DR	PRINTS; PR00453; VWFADOMAIN.
CC	-1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND	DR	SMART; SM00191; Int_alpha; 4.
CC	GRANULOCYTES.	DR	SMART; SM00327; VWA; 1.
CC	-1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS	DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC	WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.	KW	PROSITE; PS02344; VWFA; 1.
CC	-1- SIMILARITY: Belongs to the integrin alpha chain family.	KW	Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC	-1- SIMILARITY: Contains 1 VWFA domain.	FT	Signal; 3D-structure; Repeat; Magnesium; Calcium.
CC	-1- SIMILARITY: Contains 7 FG-GAP repeats.	FT	SIGNAL 1 16
CC	-1- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;	FT	CHAIN 17 1152
CC	WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".	FT	DOMAIN 17 1104
CC	-----	FT	TRANSMEM 1105 1128
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	FT	DOMAIN 1129 1152
CC	-----		
CC			INTEGRIN ALPHA-M.
			EXTRACELLULAR (POTENTIAL).
			POTENTIAL.
			CYTOPLASMIC (POTENTIAL).

between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

Query Match 99.5%; Score 5846.5; DB 1; Length 1152;									
Best Local Similarity 99.2%; Pred No. 0;									
Matches 1128; Conservative 7; Mismatches 1; Indels 1; Gaps 1;									
QY	1	FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEI	VAANQKSLYQCDYSTGSCBPI	60					
DB	17	FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEI	VAANQKSLYQCDYSTGSCBPI	76					
QY	61	RLQVPEAVNMSLGLSLAATSPOLLACGPTVHQTCSNT	VYKGLCFGLFNSLRQPOK	120					
DB	77	RLQVPEAVNMSLGLSLAATSPOLLACGPTVHQTCSNT	VYKGLCFGLFNSLRQPOK	136					
QY	121	FPFALRGCPQEDSDIAFLVDGSGSIIPHDPRAKEFI	STVMEQLKSKTFLSLMOYSEEP	180					
DB	137	FPFALRGCPQEDSDIAFLVDGSGSIIPHDPRAKEFI	STVMEQLKSKTFLSLMOYSEEP	196					
QY	181	RIHFTFKFQNNPNSRLIKPTQLLGRTHATGIRKVV	RELFNITNGARXNAFKILLI	240					
DB	197	RIHFTFKFQNNPNSRLIKPTQLLGRTHATGIRKVV	RELFNITNGARXNAFKILLI	256					
QY	241	TDGEKFGPLGYEDYVPEADREGVIRYVIGVDAPFR	SEKSRQELNTVASKPRDRHVFQIN	300					
DB	257	TDGEKFGPLGYEDYVPEADREGVIRYVIGVDAPFR	SEKSRQELNTVASKPRDRHVFQIN	316					
QY	301	NFEALKTIONQLREKIPALBGTQTSSTSSPHEMSQ	GEFSAAITNSGILLSTVGSYDWAG	360					
DB	317	NFEALKTIONQLREKIPALBGTQTSSTSSPHEMSQ	GEFSAAITNSGILLSTVGSYDWAG	376					
QY	361	GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAI	ILNRVQSLVGLGAPRYQHIGLIVANFR	420					
DB	377	GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAI	ILNRVQSLVGLGAPRYQHIGLIVANFR	436					
QY	421	QNTGWESNANYKTQIGAYPGASICSDVDVDSNGST	DLVLGAPHYQTRGGQSVVCP	480					
DB	437	QNTGWESNANYKTQIGAYPGASICSDVDVDSNGST	DLVLGAPHYQTRGGQSVVCP	496					
QY	481	PRGQARWQCDVAVLGEQCPNGRFGAALTVLGDV	NBKLTDVALGAPCEEDNRCVAVLP	540					
DB	497	PRGQARWQCDVAVLGEQCPNGRFGAALTVLGDV	NBKLTDVALGAPCEEDNRCVAVLP	555					
QY	541	HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLG	SGGDLTMDGLVBLTVGAQGHVLLLRSQ	600					
DB	556	HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLG	SGGDLTMDGLVBLTVGAQGHVLLLRSQ	615					
QY	601	PVLRVKAMEFNPREVARNVFCNDQVVKKEGSEVR	VCLHVQKSTRDLRGQIQSVVT	660					
DB	616	PVLRVKAMEFNPREVARNVFCNDQVVKKEGSEVR	VCLHVQKSTRDLRGQIQSVVT	675					
QY	661	YDLALDSGRPHSRVAFNETKNSTRQTQVLGTQCT	ETLKLQLPNCIEDPVSPIVLRNF	720					
DB	676	YDLALDSGRPHSRVAFNETKNSTRQTQVLGTQCT	ETLKLQLPNCIEDPVSPIVLRNF	735					
QY	721	SLVGTPLSAFGLNRPVLAEDAQLFTALPFFKNG	CNDNICODDLSITPFSMLDCLVVG	780					
DB	736	SLVGTPLSAFGLNRPVLAEDAQLFTALPFFKNG	CNDNICODDLSITPFSMLDCLVVG	795					
QY	781	GRPEFNVTVVRNDGSDSVRTQVTPFFPLDLSR	KVSTLQNGORSQSWELACESASSTEV	840					
DB	796	GRPEFNVTVVRNDGSDSVRTQVTPFFPLDLSR	KVSTLQNGORSQSWELACESASSTEV	855					
QY	841	SGALKSTSCSINHPIPPENSEVTFNITPDVDSK	ASLGNKLLKANVTSENNPRNTKTEP	900					
DB	856	SGALKSTSCSINHPIPPENSEVTFNITPDVDSK	ASLGNKLLKANVTSENNPRNTKTEP	915					
QY	901	QLELPVKYAVMVVTSRGVSTKYLNTASENTSRV	WQHGYQVSNLQGSRLPSLVLVVPV	960					
DB	916	QLELPVKYAVMVVTSRGVSTKYLNTASENTSRV	WQHGYQVSNLQGSRLPSLVLVVPV	975					
QY	961	RLNQTVIMDRPQVTFSENLSSCTKTKERLPSHS	DFLAELRKAPVNVCSIAVCQRIQCDIP	1020					
DB	976	RLNQTVIMDRPQVTFSENLSSCTKTKERLPSHS	DFLAELRKAPVNVCSIAVCQRIQCDIP	1035					

RESULT 2
ITAM_MOUSE
ID ITAM_MOUSE STANDARD; PRT; 1153 AA.
AC P05555;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1991 (Rel. 17, last sequence update)
DE 28-FEB-2003 (Rel. 41, last annotation update)
DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1).
GN ITGAM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX SEQUENCE FROM N.A.
RP MEDLINE=88312584; PubMed=3044779;
RA Pycela R.;
RT "Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor.";
RL EMBO J. 7:1371-1378(1988).
[2]
RP SEQUENCE OF 11-45 FROM N.A.
RC STRAIN=BALB/c; TISSUE=Spleen;
RX MEDLINE=86287312; PubMed=2942940;
RA Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E., Larson R.S., Roberts T.M., Springer T.A.;
RT "A partial genomic DNA clone for the alpha subunit of the mouse complement receptor type 3 and cellular adhesion molecule Mac-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).
[3]
RP SEQUENCE OF 17-28.
RX MEDLINE=85188276; PubMed=3887182;
RA Springer T.A., Teplow D.B., Dreyer W.J.;
RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon.";
RL Nature 314:540-542(1985).
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN MAST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED GLOMERULONEPHRITIS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION, IN RESPONSE TO A IMPAIRED DEGRANULATION AND PHAGOCYTOSIS, EVENTS THAT APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP OBESITY.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X07640; CAA30479.1; --
 EMBL; M14293; AAA39484.1; --
 PIR; S00551; S00551.
 HSSP; P11215; IABX.
 MGD; MGI:96607; Itgam.
 InterPro; IPR000413; Integrin_alpha.
 InterPro; IPR002035; WFP_A.
 Pfam; PF01839; FG-GAP_3.
 Pfam; PF00357; integrin_A; 1.
 Pfam; PF00092; vwa; 1.
 PRINTS; PR01185; INTEGRINA.
 PRINTS; PR00453; WVFADOMAIN.
 SMART; SM00191; Int_alpha; 5.
 SMART; SM00327; WVA; 1.
 PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 PROSITE; PS0234; WVFA; 1.
 Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 Signal; Calcium; Repeat.
 SIGNAL 1 16
 CHAIN 17 1153 INTEGRIN ALPHA-M.
 DOMAIN 17 1105 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 1106 1129 POTENTIAL.
 DOMAIN 1130 1153 CYTOPLASMIC (POTENTIAL).
 REPEAT 31 84 FG-GAP 1.
 REPEAT ? ? FG-GAP 2.
 DOMAIN 164 350 WVFA.
 REPEAT 337 400 FG-GAP 3.
 REPEAT 401 452 FG-GAP 4.
 REPEAT 454 515 FG-GAP 5.
 REPEAT 517 575 FG-GAP 6.
 REPEAT 580 632 FG-GAP 7.
 CA_BIND 465 473 POTENTIAL.
 CA_BIND 529 537 POTENTIAL.
 SITE 1132 1136 GPPKR MOTIF.
 DISULFID 66 73 BY SIMILARITY.
 DISULFID 105 123 BY SIMILARITY.
 DISULFID 654 711 BY SIMILARITY.
 DISULFID 770 1023 BY SIMILARITY.
 DISULFID 1028 1033 BY SIMILARITY.
 CARBOHYD 58 58 N-LINKED (GLCNAC. .)
 CARBOHYD 86 86 N-LINKED (GLCNAC. .)
 CARBOHYD 391 391 N-LINKED (GLCNAC. .)
 CARBOHYD 696 696 N-LINKED (GLCNAC. .)
 CARBOHYD 734 734 N-LINKED (GLCNAC. .)
 CARBOHYD 772 772 N-LINKED (GLCNAC. .)
 CARBOHYD 801 801 N-LINKED (GLCNAC. .)
 CARBOHYD 881 881 N-LINKED (GLCNAC. .)
 CARBOHYD 907 907 N-LINKED (GLCNAC. .)
 CARBOHYD 941 941 N-LINKED (GLCNAC. .)
 CARBOHYD 980 980 N-LINKED (GLCNAC. .)
 CARBOHYD 994 994 N-LINKED (GLCNAC. .)
 CARBOHYD 1022 1022 N-LINKED (GLCNAC. .)
 CARBOHYD 1045 1045 N-LINKED (GLCNAC. .)
 CARBOHYD 1051 1051 N-LINKED (GLCNAC. .)
 CARBOHYD 1076 1076 N-LINKED (GLCNAC. .)
 SEQUENCE 1153 AA; 127480 MW; 178DB988AECB0343 CRC64;
 Query Match 76.1%; Score 4470; DB 1; Length 1153;
 Best Local Similarity 73.9%; Pred. No. 5e-291;
 Matches 841; Conservative 146; Mismatches 149; Indels 2; Gaps 2;
 1 FNLDTEHMTFQENARGFGQSVVLOGSRVVGAPQEIIVANQRGSLYQCDYSTGSCPEI 60
 17 FNLDTEHMTFQENAKGFGQNVVLOGSGTISVVAAPQAKAVNQIGALYQCDYSTSRCHPI 76

QY 61 RIQVPEAVNMSLGLSLAATTSPOLLACQPTVHOTCSENTYVKGCLFPLFGSNLRQOPQK 120
 DB 77 PLQVPPERVNMSLGLSLAVSTVPOLLACQPTVHONCKENTYVNGLCYLFGLSNLRPQQ 136
 QY 121 PPEALRCPQSDSIAPLVDGSGSIIPHDPRAKEPISTWEOQLKSKTLPFLSMQYSEEP 180
 DB 137 PPEALRCPQSDSIVFLIDGSGSINNIDPQKKEFYSTWEOQFKSKTLPFLSMQYSEEF 196
 QY 181 RIHTEFEFQNNPNSRLIKPIFOLLGRTHATGIRKVVRELFNITNGARKNAKFIILI 240
 DB 197 RIHTEFDFKRNPSRSHVSPIKOLNERTKTSASIRKVVRELFHKTNGARENAKILVVI 256
 QY 241 TDGEKFGDPLGYEDVPEADREGVIRVYVGDVAFRSEKSRQELNVTASKPRPRHVPQIN 300
 DB 257 TDGEKFGDPLDYKDVPEADRAGVIRVYVGVGNAPNKPQSRRELDITASKPAGHFVQVD 316
 QY 301 NPEALKTIQNLREKIPIALEGTOTGSSSSPEHMSQEGFSAAITSNGLPLTGVSGYDWAG 360
 DB 317 NPEALNTIQNLQEKIPALBGTOTGSSSSPEHMSQEGFSASITSNGLPLTGVSGYDWAG 376
 QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVLGAPRYOHIGLVANFR 420
 DB 377 GAFLYTSKDKVTFINMTRVDSMDNDAYLGASAVILNRVQSLVLGAPRYOHIGLVANFR 436
 QY 421 QNTGMWESNANVGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYVETRGQSVQCP 480
 DB 437 ENPQWTEPHTSINGSGISYFGASLCSVDMDADGNTNULIGAPHYVETRGQSVQCP 496
 QY 481 PRGQSRMCCDAVLYGEOQWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAYLF 540
 DB 497 PRG-PARWQCEALLHGDQGHMGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAYLF 555
 QY 541 HGTSGGSIQSPHSQRIAGSKLSPLQYFGOSLGGQDLTMDGLVDTLTVGAQGHVLLRSQ 600
 DB 556 YGASTASLSASHSHRIIGAHFSPGLQYFGOSLGGQDLTMDGLMDLAQVGAQGHVLLRSQ 615
 QY 601 PVLRYKAJMEFNPREVARNVTECNQVWKGAGEVRVCLAVQKSTRDLREGQIQSVVT 660
 DB 616 PVLRYKAJMEFNPREVARNVTECNQVWKGAGEVRVCLAVQKSTRDLREGQIQSVVT 675
 QY 661 YDLALDSGRPHSRVNFETKSTRQTVGLTQTCETLKLQLPNCIEDPSPVILRLNF 720
 DB 676 YDLALDPVRSIRIRAFFDETKNTRTQVGLMQKCETLKLILPCVDDSDVSPVILRLNF 735
 QY 721 SLVGTPLSAGNLRLVLAEDAORLFTALPPPEKNCNDNI CODDLSITPSFMSLCLVVG 780
 DB 736 TLVGBPLRSFGNLRLPVLAEDAORFTAMPFFPKNCNDNISICODDLSITPSFMSLCLVVG 795
 QY 781 GPREFNVTVVRNDCEDSYRTQVTFPPLDLSYRKVSTLQONORSORSWL-ACESASSTE 839
 DB 796 GPQDFNMSVTLRNDGEDSYGTQVTVVYPSGLSYRKDSASQNPFTKXPMFVFPKPAESSSSE 855
 QY 840 VSGALKSTSCSINHPTFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTE 899
 DB 856 GHGALKSTTWNINHPTFPANSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTE 915
 QY 900 FQLELPVKYAVYVTVSHGVSTKYLNFTASENTSRVQVQVSNLQSRSLPISLVLP 959
 DB 916 FQLELPVKYAVYVTVSHGVSTKYLNFTASENTSRVQVQVSNLQSRSLPISLVLP 975
 QY 960 VRLNQTVIWDNRQVTFSENLSSCTHKEPLPSHSDPLAELRKAPVNVCSIAVQCORQCDI 1019
 DB 976 VQINNVTVWDHPQVTFPSQNLSSACHTEQKSPHSPRDLERTPLNCSVAVCKRLQCDL 1035
 QY 1020 PFPGIQEBFNATLKGNLSPDWIKTSHNHLIVSTAEIILFNDSVFTLLPQGAQFVRSQTE 1079
 DB 1036 PSFNTQEIFNVTLKGNLSPDWIKTSHGHLILVSSTEILFNDSAPALLPQESYVRSKTE 1095
 QY 1080 TKVEPEVNPPLIVGSSVGGILLILALITAAALKGLGPKKQYKQKMSGGPGASPO 1137
 DB 1096 TKVEPEVNPPLIVGSSIGGLVLLALITAGLYKLGFPRQYKQKMMNEAAPQADAPPO 1153

RESULT 3	
ITAX_HUMAN	
ID ITAX_HUMAN	STANDARD; PRT; 1163 AA.
AC P20702;	
DT 01-FEB-1991 (Rel. 17, Created)	
DT 01-FEB-1991 (Rel. 17, Last sequence update)	
DT 10-OCT-2003 (Rel. 42, Last annotation update)	
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95	
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).	
GN ITGAX OR CD11C.	
OS Homo sapiens (Human)	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX NCBI_TaxID=9606;	
RN [1]	
RP SEQUENCE FROM N.A.	
RX MEDLINE=8816645; PubMed=3327687;	
RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;	
RT "CDNA cloning and complete primary structure of the alpha subunit of	
RT a leukocyte adhesion glycoprotein, p150,95";	
RL EMBO J. 6:4023-4028(1987).	
RN [2]	
RP SEQUENCE FROM N.A.	
RX MEDLINE=90153906; PubMed=2303426;	
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;	
RT "Genomic structure of an integrin alpha subunit, the leukocyte	
RT p150,95 molecule.";	
RL J. Biol. Chem. 265:2782-2788(1990).	
RN [3]	
RP ERRATUM.	
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;	
RL J. Biol. Chem. 265:12750-12751(1990).	
RN [4]	
RP SEQUENCE OF 20-43.	
RX MEDLINE=87167596; PubMed=3549901;	
RA Miller L.J., Wiebe M., Springer T.A.;	
RT "Purification and alpha subunit N-terminal sequences of human Mac-1	
RT and p150,95 leukocyte adhesion proteins.";	
RL J. Immunol. 138:2381-2383(1987).	
CC -!	FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT
CC	RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL
CC	INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY
CC	IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.
CC -!	SUBUNIT: HETERO DIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X
CC	ASSOCIATES WITH BETA-2.
CC -!	SUBCELLULAR LOCATION: Type I membrane protein.
CC -!	TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC	GRANULOCYTES.
CC -!	DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC	WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!	SIMILARITY: Belongs to the integrin alpha chain family.
CC -!	SIMILARITY: Contains 1 VWFA domain.
CC -!	SIMILARITY: Contains 7 FG-GAP repeats.
CC -!	DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
CC	WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR EMBL; M81695; AA55180.1; -.	
DR EMBL; Y00093; CA68283.1; -.	
DR EMBL; M29165; -; NOT ANNOTATED CDS.	
DR EMBL; M29487; AA51620.1; ALT SEQ.	
DR EMBL; M29482; AA51620.1; JOINED.	
DR EMBL; M29483; AA51620.1; JOINED.	
DR EMBL; M29484; AA51620.1; JOINED.	
DR EMBL; M29485; AA51620.1; JOINED.	

DR EMBL; M29486; AA51620.1; JOINED.	
DR PIR; A36584; RWHUIC.	
DR PDB; 1N3Y; 18-FEB-03.	
DR Genew; HGNC:6152; ITGAX.	
DR MIM; 151510;	
DR GO; GO:0008305; C:integrin complex; TAS.	
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.	
DR GO; GO:0004872; F:receptor activity; TAS.	
DR GO; GO:0007155; P:cell adhesion; TAS.	
DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.	
DR InterPro; IPR000413; Integrin_alpha.	
DR InterPro; IPR002035; VWFA.	
DR Pfam; PF01839; FG-GAP; 3.	
DR Pfam; PF00357; Integrin_A; 1.	
DR Pfam; PF00092; vwa; 1.	
DR PRINTS; PR01185; INTEGRINA.	
DR PRINTS; PR00453; VWFADOMAIN.	
DR SMART; SM00191; Int_alpha; 5.	
DR SMART; SM00327; VWFA; 1.	
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.	
DR PROSITE; PS00234; VWFA; 1.	
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;	
KW Signal; Magnesium; Calcium; Repeat; 3D-structure.	
FT SIGNAL 1 19	
FT CHAIN 20 1163	INTEGRIN ALPHA-X.
FT DOMAIN 20 1107	EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1108 1128	POTENTIAL.
FT DOMAIN 1129 1163	CYTOPLASMIC (POTENTIAL).
FT REPEAT 34 87	FG-GAP 1.
FT REPEAT ? ?	FG-GAP 2.
FT DOMAIN 165 351	VWFA.
FT REPEAT ? ?	FG-GAP 3.
FT REPEAT 402 453	FG-GAP 4.
FT REPEAT 455 517	FG-GAP 5.
FT REPEAT 518 576	FG-GAP 6.
FT REPEAT 581 633	FG-GAP 7.
FT CA BIND 466 474	POTENTIAL.
FT CA BIND 530 538	POTENTIAL.
FT CA BIND 593 601	POTENTIAL.
FT SITE 1131 1135	GFGR MOTIF.
FT DISULFID 69 76	BY SIMILARITY.
FT DISULFID 108 126	BY SIMILARITY.
FT DISULFID 655 712	BY SIMILARITY.
FT DISULFID 771 777	BY SIMILARITY.
FT DISULFID 848 863	BY SIMILARITY.
FT DISULFID 998 1022	BY SIMILARITY.
FT DISULFID 1027 1032	BY SIMILARITY.
FT CARBOHYD 61 61	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 89 89	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 697 697	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 735 735	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 899 899	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 939 939	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1050 1050	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 490 490	G -> A (IN REF. 2).
FT CONFLICT 756 756	L -> D (IN REF. 2).
SQ SEQUENCE 1163 AA; 127885 MW; 6C4E19CC3F62A473 CRC64;	
Query Match	59.0%; Score 3469; DB 1; Length 1163;
Best Local Similarity	61.0%; Pred. No 5.2e-224;
Matches 689; Conservative 141; Mismatches 293; Indels 6; Gaps 4;	
Qy 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVYWGAPQEIIVANQSGSLYQCDYSTGSCPEI 60	
Db 20 FNLDTBELTAFRVDSAGFGDSVVQYANSVVWGAPQKITAANQTGGLYQCGYSTGACEPI 79	
Qy 61 RLOVPVENVNMISGLSIAATTSPPQLACGPTVHOTCSENTYVKGLCFPLGSLNRQOPOK 120	
Db 80 GLQVPPPEAVNMISGLSLASTTSBSQLACGPTVHCEGRNMYTGLCFLLPPT--QUTQR 137	
Qy 121 FPEALRCGPQEDSDIAFLVDGSGSIIPHDFRAKEFIISTVMEQLKSKSLFSLMQYSEEF 180	

138 LPVSRQECPRQEQDIIVFLIDGSSISRRNFATMNFVRAVISOQFQSPQPSLMQFSNKP 197
 181 RIHTFKPQNNPNPSLKIPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
 198 QHTFTFEERFRTNPLSLLASVHQLQGTFTYATAIQNVVHRLFHASYGARRDATKILIVI 257
 241 TDGEKFGDPLGYEDVPEADREGVIRYVGVGDAPFSEKSRQBLNTVASKPRDHFQJIN 300
 258 TDGKKGDSLDYKDVIPMDAAGIIRIYAGVGLAFQNRNSWKEINDIASKPSQEHFKYE 317
 301 NFPAKTIQNLREKIFAIEGTOTGSSSFEHMSQEGFSAALTSNGPILLSVTGSDVWAG 360
 318 DFDALXDIQNLREKIFAIEGTOTGSSSFEHMSQEGFSAALTSNGPILLSVTGSDVWAG 377
 361 GVFLYTSKSKSTINRVDSDMNDAYLGAARAIILNRVQSILVIGAPRYOHLGLVAMER 420
 378 GAFLYPNNMPTFINMSQNVDRDLSGLYSTELALWKGVSQSLVIGAPRYOHTGRAVIFT 437
 421 QNTGMWESANVKGTOIGAFYGASLCSVDVDSNGSDTLVLIGAPHYYEYTRGQVSVCSPL 480
 438 QVSRQWRKAEVGTQIGSYFGASLCSVDVDTGSDTLVLIGAPHYYEYTRGQVSVCSPL 497
 481 PRQARWQCDAYLYGEGQWGRFGAALTVLGDVNGDKLTDVAIGARCEENRGAVILF 540
 498 PRQWR-RWKCDAVLYGEGQWGRFGAALTVLGDVNGDKLTDVIGAPGEBENRGAVILF 556
 541 HGTSGGSGIPSHSORIAGSKLSPLQYFGQSLGGQDLTMDGLVDLTGAGQVHLLRSQ 600
 557 HGVLGSPISPSHGORIAGSKLSPLQYFGQSLGGQDLTMDGLVDLTGAGQVHLLRSQ 616
 601 PVLRVKAEMFNPREVARNVFQNDQWKEAGEVRLVHVKSTRDLRREGQISQSVT 660
 617 PVLVWGVSGQFIPAEIPRFAFECEQVSEQTLVQSNICLYIDKRSKNLGSRLDQSSVT 676
 661 YDLALDSGRPHSAVNETKNSRTOVGLTQTCETLKLQNPNCIEPVPSPVILRLNF 720
 677 LDIALDPGRSLPRATQSTQKRSLSRVVLGKAHCENFNLLPSCVEDSVPTILRLNF 736
 721 SLVGTPLSAPNLRPLAEDAQELFTALPPFERNKGNMDCDDLSITFSFMSLDCLVVG 780
 737 TLVGKPLAPNLRPLAEDAQELFTALPPFERNKGNMDCDDLSITFSFMSLDCLVVG 796
 781 GPREENVTVVRDGDSDYRTOVTFPFLDLSVRKYSTLQNDQSRQSWLACASSTEV 840
 797 SNLEMAVWVNDGSDSYRTOVTFPFLDLSVRKYSTLQNDQSRQSWLACASSTEV 854
 841 SGALKSTCSINHPITPENSEVTFTFVDVSKASLGNKLLKANTYSENMPRYNKTFF 900
 855 SQGTWSTSCRINHLIFRGAQITFLATFVSPRAVLGDRLLLTANVSSENNTPRTSKTF 914
 901 QLELPKYAVYVWYSHGYSYKYLNTAS-ENTSRVWQHOYQVNLGQSLPLSLVFLYP 959
 915 QLELPKYAVYVWYSHGYSYKYLNTAS-ENTSRVWQHOYQVNLGQSLPLSLVFLYP 974
 960 VRLNQTVINDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDI 1019
 975 VELNQEAVMWVESHVSPQNPPLSCSEKIAFPASDFLAHQKPNVLDGSLAGCLRPRCDV 1034
 1020 PFGIOBEFNATIKHLSFNYKTSNHLIIVSTAEILNDVFTLLPQCGAFVRSOTE 1079
 1035 PPSVQBEELFTKXNLSFGVWVQFQILQKRVVSVVAEITPDTYSVSQLPQGFAPMAQT 1094
 1080 TKVEPFEVNPFLPIVGVSSVGGLLALLALITALYALYKLGFFKRYQKMWSE 1128
 1095 TVLEKYKVNPTPLIVGVSSVGGLLALLALITALYALYKLGFFKRYQKMWSE 1143

RESULT 4
 ITAD HUMAN
 AC Q13379; Q15575; Q15576;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)

28-PEB-2003 (Rel. 41, Last annotation update)
 Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADBE2).
 ITGAD.
 Homo sapiens (Human).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 [1] SEQUENCE FROM N.A.
 TISSUE=Splice;
 MEDLINE=95111956; PubMed=8777714;
 Van der Vliet M., Le Trong H., Wood C.L., Moore P.F., St John T.,
 Staunton D.E., Gallatin W.M.;
 "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-
 3";
 Immunity 3:683-690 (1995).
 [2] SEQUENCE OF 1-235 FROM N.A.
 MEDLINE=20187620; PubMed=10722744;
 Noti J.D., Johnson A.K., Dillon J.D.;
 "Structural and functional characterization of the leukocyte integrin
 gene CD11d. Essential role of S₁ and S₃";
 J. Biol. Chem. 275:8959-8969 (2000).
 [3] SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
 MEDLINE=96257236; PubMed=8666289;
 Wong D.A., Davis E.M., LeBeau M., Springer T.A.;
 "Cloning and chromosomal localization of a novel gene encoding a human
 beta 2-integrin alpha subunit";
 Gene 171:291-294 (1996).
 [4] INTERACTION WITH VCAM1.
 MEDLINE=99059842; PubMed=9841932;
 Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W.,
 Hoffman P.A., Staunton D.E., Bochner B.S.;
 "alpha₂beta₂ integrin is expressed on human eosinophils and functions
 as an alternative ligand for vascular cell adhesion molecule 1
 (VCAM-1)";
 J. Exp. Med. 188:2187-2191 (1998).
 [5] INTERACTION WITH VCAM1.
 MEDLINE=99370002; PubMed=10438935;
 Van der Vliet M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,
 Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
 "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a
 binding interface between I domain and VCAM-1";
 J. Immunol. 163:1984-1990 (1999).
 CC - FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND
 CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-
 BORNE PATHOGENS, PARTICULATE MATTER, AND SENESECENT ERYTHROCYTES
 FROM THE BLOOD.
 CC - SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D
 ASSOCIATES WITH BETA-2.
 CC - SUBCELLULAR LOCATION: Type I membrane protein.
 CC - TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL
 LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON
 TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN
 ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.
 CC - DOMAIN: THE INTEGRIN I-DOMAIN (INSET) IS A VFMA DOMAIN. INTEGRINS
 WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC - SIMILARITY: Belongs to the integrin alpha chain family.
 CC - SIMILARITY: Contains 1 VFMA domain.
 CC - SIMILARITY: Contains 7 FG-GAP repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 or send an email to license@isb-sib.ch.


```

FT CARBOHYD 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 649 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 670 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 726 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 730 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 862 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 885 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 897 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1060 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 954 Q -> QCVHGVEMQTSKQILCRPAGDAEHVQSGELPC
      PWCSEAFNRINRAGPCR (in isoform 2).
      /FTID=VSPF 902738.
      R -> W (IN REF. 1 AND 2).
      Y -> I (IN REF. 2).

FT CONFLICT 214
FT CONFLICT 660
FT STRAND 155
FT STRAND 164
FT TURN 165
FT HELIX 169
FT TURN 186
FT STRAND 191
FT STRAND 202
FT STRAND 208
FT HELIX 213
FT HELIX 217
FT TURN 222
FT STRAND 229
FT HELIX 233
FT TURN 244
FT HELIX 247
FT TURN 250
FT TURN 253
FT STRAND 256
FT HELIX 274
FT TURN 277
FT STRAND 280
FT HELIX 288
FT HELIX 293
FT TURN 298
FT HELIX 300
FT HELIX 307
FT TURN 311
FT TURN 318
FT HELIX 319
FT TURN 329
SQ SEQUENCE 1170 AA; 128819 MW; 39A7AF92EF286FC0 CRC64;

Query Match 26.5%; Score 1557.5; DB 1; Length 1170;
Best Local Similarity 34.7%; Pred. No. 4.9e-96;
Matches 407; Conservative 206; Mismatches 460; Indels 101; Gaps 37;

QY 1 FNLDPENAMTFO--ENARGFGQSVVQLQSGSRVVVGAPQISVAANQSGLYQCDYSTGSC 58
DB 26 YNLDVVGARSFSPRAGRHFGYVLQV--GNGVIVGAPGE---GNSGSLYQCSQSTGCL 81
QY 59 PIRLVQPVVEANMISGLSLAATSPQILLACGPTVHQTCSNTYVKGICFLFGNLR--- 115
DB 82 PVTLR--GSNYTSKYLGMTLATDPTDGSILACDGLSRTCDQNTYLSGLCYLFRONLQGP 140
QY 116 -QQQKPPFALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKGFISTVMEOLKSKTFLSLM 174
DB 141 LQGRGFCQECIKG-----NVDLVFLFDGSMISQDPDEFQKILDFMKDVMKLSNTSYQPAV 196
QY 175 QYSEBFRHFTKFEQNNPNRSLIKPITQLGRTHATGIRKVVVRELFINITNGARKNAP 234
DB 197 QPSTSYKTFDFSDYVKRQDPALLKHVKHMLLTNTFGAINVATVFEVBELGARDPAT 256
QY 235 KLLIITDGEKFGDPLGYEDVTPEDREGVIRYVLGVGDAPRSEKSRQELNATVASKPRD 294
DB 257 KVLIIITDGE--ATDSGNIDAKD-----IIRYIIGIKGHFQTKESQETLHKFASKPASE 309
QY 295 FVFQINNFEALXKIQNLKRIKFAIEGTQGTGSSSSSEHMSQEGFSAAITNGPLLSTVG 354

```

```

DB 310 FVKILDTFEKLDLFTLQKIYIEGTSKODLTSFNMELSSGISADLGRHVAWGAVG 369
QY 355 SYDWAGGVF--LYTSKSKSTFINMTRVDSMDNDAYLGYAAA--IILRNEVQSLVILGAPRYQH 412
DB 370 AKDWAGFLDKADLQDDTFIGNELTPEVRAGYLGVTVTWLPSPRQKTSLLASGAPRYQH 429
QY 413 IGLVAMPR--QNTGMWESNANVIGTGIGAFGASLCSVDVDSNGSTDLVILGAPHYEQT 470
DB 430 MGRVLLFQBPQGGHWSQVQTIHGTOIGSYFGGELCGVDVDQDGETELLIGLAPLFYGEQ 489
QY 471 RGGVSVCPPLRGORARWQCDAV--LVGEQGGPWGRGAALTVLGDVNGDKLTDAVLCAP 528
DB 490 RGRVFIY-----QRRQLGFEFEVSELQDPCYPLGRFGEAITALTDINGDLVDVAVGAP 544
QY 529 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGSQDQDLMGLVDLTV 588
DB 545 LEE--QGAVYIFNGRHG--GLSPQPSQRIEQTQVLSGIGQWFGFRSITHGVKIDLEDGLADVAV 601
QY 589 GAQGHVLLFSQPVLRVKATMEFNPREVARNFECNDQVV--KGKEAGEVAVVCLHVQKSTR 647
DB 602 GAEQMIVLSSRPVDMVTLMSFSPABI PVHVECSYSTSNKKEGVNITICQI--KSLY 660
QY 648 DRLREGIQSVVYDYLALDSGRPHSRAVFNETHNSTRQTOVLGLTQTCETLKLQLPNCI 707
DB 661 PQF--QGRLVANLTVTLQDGHRTERRRGLFPGGRHELRNIAVT--TSMSCDTSFHFVFCV 718
QY 708 EDPVSPVILNPSL---VGTPLS--AFGN-----LRPLAEDAQLRFTALFPPEKNCN 757
DB 719 QDLISPINVSLNFSLMEEBETPRDQRAQKDIPIILRPSLHSETWEI-----PPEKNGE 773
QY 758 DNICQDDLSITFMSLDCLVVGPRFNVTVTRNDGDSYRTQVTFPPPLDLISYRKVS 817
DB 774 DKXCEANLRVSPARSRALRLTAFASLSVELSLNLEEDAYVYQLDHLHFPGLSPKRV 833
QY 818 TLQNRQSRWRPLACES--ASSTEVSGALSKTSCSINHPIPEPENSEVTFNITPDVDSKAS 875
DB 834 ML---KPHSQIPVSCBELPESRLLSRAL---SCNVSSPIFKAGHSVALQOMFNLTNVS 887
QY 876 LGNKLLKANVTSENN---MPTNTEFQLELPVKYAVTVVTVTSHGVSTKYNFTASBN 931
DB 888 WEDSVELHANVTNNEDSDILEDNASATTI---IPILYINILIQDQDSDTLYVSFTPKGP 944
QY 932 TSRVMQHOYQV---SNLQGRSLP--ISLVFLVPLVRLNQTVDNRPOVTFSENLSSTCHTK- 986
DB 945 KIHQVKHMYQVRIQPSIHDHNIPTLEAVGVGPQPPSEGGPIITHQMSVQMEPPV--PCHYED 1002
QY 987 -ERLPSSHSD--FLAELKAPVWNCSTAVCQRIQCDIPFGIQEENFATLKNLSFDMYIK 1043
DB 1003 LERLPDRAEPCLPALFRCPVV-----FRQELVQVIGTLELVGEIE 1044
QY 1044 TSNHLLIVSTABILFNDVSFTLLPGQGAFFVRQGTETKVPFPFVFNPLPLIVGSSVGLL 1103
DB 1045 AS-SMPSLCSLSISFNSKHFHLYGSNASL-AQVVMKVDVYVEKQMLYLYLVLSGIGLL 1102
QY 1104 LLALITAAALKYKAGFKKQYKDMMSG--GPPGAP 1136
DB 1103 LLLIFIVLYKVGFFKRNLRKEAGRGVNGIP 1136

RESULT 6
ITAL MOUSE STANDARD; PRT; 1163 AA.
ID ITAL MOUSE
AC P24063,
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1 alpha chain) (Leukocyte function associated molecule 1, alpha chain) (CD11a).
DE ITGAL OR LFA-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1] _SEQUENCE FROM N.A.
 RP MEDLINE=91268576; PubMed=2051027;
 RX MEDLINE=91268576; PubMed=2051027;
 RA Kaufmann Y., Tseng E., Springer T.A.;
 RT "Cloning of the murine lymphocyte function-associated molecule-1
 alpha-subunit and its expression in COS cells.";
 RL J. Immunol. 147:369-374(1991).
 RN [2]
 RP SEQUENCE OF 24-42.
 RX MEDLINE=85188276; PubMed=3887182;
 RA Springer T.A., Teplow D.B., Dreyer W.J.;
 RT "Sequence homology of the IFA-1 and Mac-1 leukocyte adhesion
 glycoproteins and unexpected relation to leukocyte interferon.";
 RL Nature 314:540-542(1985).
 CC !- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
 CC ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA,
 CC INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL
 CC MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES
 CC AND MONOCYTES. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-L
 CC SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED
 CC LEUKOCYTE RECRUITMENT.
 CC !- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L
 CC ASSOCIATES WITH BETA-2.
 CC !- SUBCELLULAR LOCATION: Type I membrane protein.
 CC !- TISSUE SPECIFICITY: LEUKOCYTES.
 CC !- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A WFPA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC !- SIMILARITY: Belongs to the integrin alpha chain family.
 CC !- SIMILARITY: Contains 1 WFPA domain.
 CC !- SIMILARITY: Contains 7 FG-GAP repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M60778; AAA39426.1; -.
 DR PIR; I56126; I56126.
 DR HSSP; P20701; ILFA.
 DR MGD; MGI:96606; Itgal.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; WVF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; WVFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; WVA_1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS50234; WVF; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Magnesium; Calcium;
 KW Repeat.
 FT SIGNAL 1 23
 FT CHAIN 24 1163
 FT DOMAIN 24 1084
 FT TRANSMEM 1085 1108
 FT DOMAIN 1109 1163
 FT REPEAT 39 88
 FT REPEAT 7 7
 FT DOMAIN 148 334
 FT REPEAT 2 2
 FT REPEAT 399 454
 FT REPEAT 455 514
 FT REPEAT 516 573
 FT REPEAT 576 628
 FT CA_BIND 466 474

QY	CA_BIND	528	536	POTENTIAL.
FT	CA_BIND	528	536	POTENTIAL.
FT	CA_BIND	588	596	POTENTIAL.
FT	SITE	1111	1115	GFPR MOTIF.
FT	DISULFID	70	77	BY SIMILARITY.
FT	DISULFID	108	126	BY SIMILARITY.
FT	DISULFID	147	159	BY SIMILARITY.
FT	DISULFID	651	705	BY SIMILARITY.
FT	DISULFID	767	773	BY SIMILARITY.
FT	DISULFID	840	856	BY SIMILARITY.
FT	DISULFID	993	1009	BY SIMILARITY.
FT	DISULFID	1017	1048	BY SIMILARITY.
FT	CARBOHYD	86	86	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	185	185	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	270	270	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	444	444	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	668	668	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	696	696	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	724	724	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	728	728	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	776	776	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	857	857	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	880	880	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	890	890	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	899	899	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	927	927	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1056	1056	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1163	1163	N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQ	SEQUENCE	1163	AA; 128343	MW; A7A307848988232F CRC64;

Query Match 26.2%; Score 1538.5; DB 1; Length 1163;
 Best Local Similarity 34.2%; Pred. No. 9.1e-95;
 Matches 401; Conservative 214; Mismatches 458; Indels 99; Gaps 36;

QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
QY	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598
DB	1	24	60	80	120	139	180	199	239	259	292	305	352	365	410	425	468	485	528	542	598</

Db 599 VAGRVRVLSRPVVDVVTLSFSPBEPVHEVECSASAREBQKHGVLKACAFRIPLT 658
Qy 647 RDLREGIQSVVTVYDLALDSCRPHSAVFNETKRNSTRQVGLTQTCTBLKQLPNC 706
Db 659 PQ--FOGRLLANLSTYQLDCHRRSRGLPFGGSHLSGNTSITP-DKSLCDFHFPIC 715
Qy 707 LEDPSPVILNLSLV---GTPLSAFGN-LRPVLAEDAQLFALPFPKNGCNDNIQC 762
Db 716 IODLSPINVLNLSLEBEGTPRQKRAMQILPISHTV-TKEIPFKNGEDKKCE 774
Qy 763 DDLSTTFMSLDCLVWGP-----REFNTVTVYRNDGSDSYRTQVFPFPLDLSYRKV 816
Db 775 ANLTSSPARS-----GPLRLMSSASLAVENTLNSGSDAYVWRLDLPFRLGSPRKV 827
Qy 817 STLQNRQQRWRNLACESASTVSGAL-KSTSCSINHPIPFENSEVTFNITFDVDSKAS 875
Db 828 EMLQ---PHSRMPVSCEBL--TEGSLSTKLKCNVSPFPKAOEVSLOQWENTLLNS 882
Qy 876 LGKLLKLLKANTSEN-NMPTNKTEFQELPVPKAVYVVTSHGVSTKYLNFTASENTSR 934
Db 883 WEDFVLNLTVCENENSLQEDNSAATHIPVLPVNLTKQENSTLYISFTPKGPKTQ 942
Qy 935 VQHOYQVSNLQGRSLPISLVPLVRLNQVIMDRPO-----VTFSENLS--TCHTK 986
Db 943 QVQHYVQV-----RIQPSAYDHNMPT-LEALVGVPRPHSEDLITVTSVQTDPLVTCHSE 996
Qy 987 E-RLPSSHDFLAEKKAFAVNCISIAVCORICDIPFGIOGBFNATLKNLSFDWYIKTS 1045
Db 997 DKRPSE---ABQCPLGV-----QFRCPVFE---RWEILLQVGTVELSKELKAS 1042
Qy 1046 HNHLIVSTABILFNDSTVFTLLPGGAPFVRQSTETKVPFEPVNPPLPLVSGVGLLLL 1105
Db 1043 -STLSCLSSLSVSNFNSKHFLYGSKA-SEAQVLVKVDLIHEKMLVYVLSGIGVLVL 1100
Qy 1106 ALITAAVKLGFYKQYKDMV-SEGPPGCAEP 1136
Db 1101 FLIFLALYKVGFFKRLKEMADGGVNGSP 1132

RESULT 7
ITAE MOUSE
ID ITAE MOUSE STANDARD; PRT; 1167 AA.
AC Q60677;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Integrin alpha M290).
GN ITGAE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC STRAIN=AKR;
RX MEDLINE=95187992; PubMed=7982170;
RA Smith T.J., Ducharme L.A., Shaw S.K., Parker C.M., Brenner M.B.,
RA Kilshaw P.J., Weis J.H.;
RT Immunity 1:393-403 (1994).
RL -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
CC CELL MONOLAYERS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-E
CC SUBUNIT GENE EXHIBIT A MARKED REDUCTION IN THE NUMBERS OF
CC INTRAEPITHELIAL LYMPHOCYTES IN THE GUT AND IN THE DEVELOPMENT OF
CC GUT-ASSOCIATED LYMPHOID AGGREGATES, SUPPORTING A SPECIFIC ROLE FOR
CC THIS INTEGRIN IN MEDIATING RETENTION OF LYMPHOCYTES IN THE
CC INTESTINAL WALL.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
CC -!- SUBCELLULAR LOCATION: type I membrane protein.

CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 VWFA repeats.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U12236; AAC52142.1; -.
CC HSP; P11215; IABX.
CC MGD; NGI:1298377; Itgae.
CC InterPro; IPR000413; integrin_alpha.
CC InterPro; IPR002035; VWF_A.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC PRINTS; PR01185; INTEGRINA.
CC PRINTS; PR00453; VWFADOMAIN.
CC SMART; SM00191; int_alpha; 3.
CC SMART; SM00327; VWA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS00234; VWFA; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Signal; Repeat; Magnesium;
CC Calcium.
CC SIGNAL 1 19 BY SIMILARITY.
CC CHAIN 20 1167 INTEGRIN ALPHA-E.
CC CHAIN 20 181 INTEGRIN ALPHA-E LIGHT CHAIN.
CC CHAIN 183 1167 INTEGRIN ALPHA-E HEAVY CHAIN.
CC DOMAIN 20 1114 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 1115 1137 POTENTIAL.
CC DOMAIN 1138 1167 CYTOPLASMIC (POTENTIAL).
CC REPEAT ? ? FG-GAP 1.
CC REPEAT ? ? FG-GAP 2.
CC DOMAIN 149 192 X-DOMAIN (EXTRA DOMAIN).
CC DOMAIN 193 384 VWFA.
CC REPEAT ? ? FG-GAP 3.
CC REPEAT 449 501 FG-GAP 4.
CC REPEAT 503 564 FG-GAP 5.
CC REPEAT 566 631 FG-GAP 6.
CC REPEAT 634 686 FG-GAP 7.
CC CA_BIND 514 522 POTENTIAL.
CC CA_BIND 578 586 POTENTIAL.
CC CA_BIND 646 654 POTENTIAL.
CC DOMAIN 185 191 GLU-RICH (ACIDIC).
CC SITE 1140 1144 GFGR MOTIF.
CC DISULFID 72 83 BY SIMILARITY.
CC DISULFID 130 164 BY SIMILARITY.
CC DISULFID 698 754 BY SIMILARITY.
CC DISULFID 814 820 BY SIMILARITY.
CC DISULFID 884 898 BY SIMILARITY.
CC DISULFID 998 1023 BY SIMILARITY.
CC DISULFID 1031 1047 BY SIMILARITY.
CC CARBOHYD 51 51 N-LINKED (GLCNAC. .)
CC CARBOHYD 256 256 N-LINKED (GLCNAC. .)
CC CARBOHYD 314 314 N-LINKED (GLCNAC. .)
CC CARBOHYD 341 341 N-LINKED (GLCNAC. .)
CC CARBOHYD 364 364 N-LINKED (GLCNAC. .)
CC CARBOHYD 418 418 N-LINKED (GLCNAC. .)
CC CARBOHYD 437 437 N-LINKED (GLCNAC. .)
CC CARBOHYD 718 718 N-LINKED (GLCNAC. .)
CC CARBOHYD 773 773 N-LINKED (GLCNAC. .)
CC CARBOHYD 829 829 N-LINKED (GLCNAC. .)
CC CARBOHYD 846 846 N-LINKED (GLCNAC. .)
CC CARBOHYD 911 911 N-LINKED (GLCNAC. .)
CC CARBOHYD 925 925 N-LINKED (GLCNAC. .)
CC CARBOHYD 968 968 N-LINKED (GLCNAC. .)

FT CARBOHYD 1013 1013 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1086 1086 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1167 AA; 128983 MW; B8331C115DCCFFD CRC64;
Query Match 19.8%; Score 1161.5; DB 1; Length 1167;
Best Local Similarity 29.0%; Pred. No 1.6e-69;
Matches 357; Conservative 214; Mismatches 457; Indels 205; Gaps 43;
QY 1 FNLDTENA--MTQENARFGQSVVQLOGRVVGAPQBIIVANORGS-----LYQCDY 52
DB 20 FNMVDWAWYATQPCAPAVLSLLHDPN-----NOTCLLVARSSNNTAALYRCAI 74
QY 53 STGSCPIRLOVPEAVNMSLGLSLAATT--SPQILLAC-GPTVHGTSCENTYVGLCLFL 109
DB 75 SI-SPEIACQ-PVEHICMPKRYQGVTLVGNHNGVLCVIOQVAKFRSLNSELGTGACSL 132
QY 110 FGSNLRQQQKPFPEALRG-----C-----POE 131
DB 133 LTPNLDLOAQAVFSDLEGFLDPCAHVDSGDYCRSGKSGTGEKKSAARRRTVEEDED 192
QY 132 DSIAFLVDGSGIIPHDPRAKEFTSTWEQL--KSKTSLFSLMOYSEEPRIHPTKEF 189
DB 193 GTHIAIVLDGSGIGFSDFOKAKNFITMWRNFYKFCFCNFALVOYGAIVOTEPDLOES 252
QY 190 QNNPNRSLIKPTQLIGRTHATGIRKVVRELFNITNGARKNAFKILILITDGEKFGDP 249
DB 253 RDINASLAKVQSVQVKEVKTASAMQHLVDNIFIPSRGSRKALKVMVLTGDIKFGDP 312
QY 250 LGYEDVPEADRGVRYVLCGDARSEKSRRELATVASKPRDHFVQINNEPEALKTIQ 309
DB 313 LNLTTVINSKMGQVRFVFAIGVGRFKNNTYRELKLIASDPRKENTFKVTNYSALDGLL 372
QY 310 NQLEKIFALGEGTQSSSFHEMSQBSAITSNGP--LLSTGVSYDWAGGVFLY--TS 367
DB 373 SKLQORIVNEGTT--VGDLALQVLAQTGFSAQILDKGVLLCTVGFANWGGALLYSTQ 429
QY 368 KEKSTFINMT-RVDS-DMNDAYLGIAAAIILNRVQSLVIGARYQHIGLVAMFRONTGM 425
DB 430 NGRGCFINQAKEDSRVQVSYLGSLAVLKHAGISYVAGAPRRKLRGAVFLRKEDR- 488
QY 426 WESNA---NVKGTOIGAYFCASLCSDVDNSGSTDVLICAPHYEYTRGQGVSVCLPR 482
DB 489 -BEDAFVRIEGEQMGYFVSCLPVDIDMDGTTDFLLVAAPYHIRGEGRVYVQVPE 547
QY 483 GQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDAIGAP-----GRENRGA 536
DB 548 -QDASPSLAHTLSGHPGLTNSRPGFAMAAGVDINQDKFTDVAIGAPLEGFGAGDSYGS 606
QY 537 VYLFHGTSGSGISPSHSORAGSKLSPLOYFGOSLSGGDLTMDGLVLTGVAQGHVLL 596
DB 607 VYIINGHSG-GLYDPSQQLIRASSVAGSLHYFGVSGGLDFNGDGLADITVSRDSAVV 655
QY 597 LRSQVILRVKAIMFNPREVARNVFECNDQVVKGEAGYRVCLHVQKS---TRDLREG 653
DB 666 LRSRPVVDLTVSMFTF-----DALPMVFIKGM--DVNLCPVDSSVVASEPGLREM 715
QY 654 QIQSVVYDLDALDSGRPHRAVFNKSTNRQTVLGLTQC-----696
DB 716 FLNFTVDV-----TKQRQLOCEDSGGQSLRKKNWGSFICEHPWLI 760
QY 697 ETEKLQLPNCTEDPVSPIVLKLNFLSLVGTPLSAFGLNR-----PVLARDQRLFTALP--P 750
DB 761 STEEL-----CEDDCSNITIKVYE-----PQTSGGRDYENPTL--DHYKEPSAIFQLP 809
QY 751 FEKVCNDNIQDDLSITFFSMGLDCLVWGSPRENTVTVRNDGDSYRTQVTFPPLD 810
DB 810 YEKDCKRVFCIAIQTLTN--ISQQLVWGVTKVETVNI SLTNSGDSYMTNMAIYFN 868
QY 811 LSVKSVTLQNRORSRWLRACBSASSTEVSGALKSTSCSNHPIFFENSEVFNITFDV 870
DB 869 LQFKKI-----QKPSDPVQCDPKPV---ASVLVWNCIKIGHFIL-KRSVNVSVTWQL 918

QY 871 DSKASLGNKLLKXANVTSENNPRTNTEFOLELPVKYAVVMVVTSHGVSTKYLNTASE 930
DB 919 BESVFPNRTADITVTISNKSLEARETR---SIQFHAFIAVLRS--PSWVYWN--TSQ 971
QY 931 NTSVMQHQYQVSNLQGRSLPISLVFLVPLNQTVMWRPQVTFSENLST-----CHT 985
DB 972 SPDSKKEPFFNVHGENLFGAVFQLQICVPIKLODF-----QIVRVKMLTKFDHTECTQ 1025
QY 986 KERLPSPSDFLAELRKAPVNCSTAVCQRIQCDIPFGCIQEEFNATLKNLSFDWYIKTS 1045
DB 1026 SQEPACGSDPVQVHKEHSHVVCAL-----TSNKENTVAAEISVVG 1065
QY 1046 FNHLILVSTA-----EILFNDVSFTLLPQOGAFVRSQETKVPFF-----EVPNPLPIIV 1095
DB 1066 HTKQLLRDVSBLPILGEISFNKSLYEGLNAE-----NHRTKITVIFLKEBETRSLPIII 1119
QY 1096 GSSVGLLILLIITALYKLGFKFRQYKDMKSE 1128
DB 1120 GSSIGGLLVVIFAILFKCGFRKRYQOOLNLE 1152
RESULT 8
ITAE HUMAN
ID ITAE_HUMAN STANDARD; PRT: 1179 AA.
AC P8570; Q9NZU9;
DT 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1
DE antigen) (CD103 antigen) (Integrin alpha-IEL).
GN ITGA8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.
RC TISSUE=Leukemia, and Lymphocytes;
RX MEDLINE=94164962; PubMed=8119947;
RA Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B.,
RA Parker C.M.;
RT "Molecular cloning of the human mucosal lymphocyte integrin alpha E
RT subunit. Unusual structure and restricted RNA distribution.";
RL J. Biol. Chem. 269:6016-6025(1994).
RN [2]
RP REVISIONS TO 88-114.
RA Parker C.M.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE OF 53-1179 FROM N.A.
RC TISSUE=Petal kidney;
RX MEDLINE=20138496; PubMed=10673275;
RA Touchman J.W., Ankster Y., Dietrich N.L., Maduro V.V., McDowell G.,
RA Shotelersuk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A.,
RA Green E.D.;
RT "The genomic region encompassing the nephropathic cystinosis gene
RT (CTNS): complete sequencing of a 200-kb segment and discovery of a
RT novel gene within the common cystinosis-causing deletion.";
RL Genome Res. 10:165-173(2000).
RN [4]
RP MUTAGENESIS OF ASP-109 AND PHE-316.
RX MEDLINE=20400502; PubMed=10837471;
RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y.,
RA Brenner M.B.;
RT "The role of alpha and beta chains in ligand recognition by beta 7
RT integrins.";
RL J. Biol. Chem. 275:25652-25664(2000).
CC -I- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
CC CELL MONOLAYERS.
CC -I- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.

[illegible]

DR	GO:	GO:0005518; F:collagen binding; TAS.	
DR	GO:	GO:0007160; P:cell-matrix adhesion; TAS.	
DR	GO:	GO:0007517; P:muscle development; TAS.	
DR	InterPro:	IPR000413; Integrin_alpha.	
DR	InterPro:	IPR002035; WVF_A.	
DR	Pfam:	PF01839; FG-GAP 3.	
DR	Pfam:	PF00092; vwa; 1.	
DR	PRINTS:	PRO1185; INTEGRINA.	
DR	PRINTS:	PRO0453; WVFADOMAIN.	
DR	SMART:	SM00191; Int_alpha; 5.	
DR	SMART:	SM00327; WVF; 1.	
DR	PROSITE:	PS00242; INTEGRIN_ALPHA; FALSE_NEG.	
DR	PROSITE:	PS0234; WVF; 1.	
KW	Integrin;	Cell adhesion; Receptor; Glycoprotein; Transmembrane;	
KW	Signal;	Repeat; Calcium; Magnesium; Polymorphism.	
FT	SIGNAL	1 22	POTENTIAL.
FT	CHAIN	23 1189	INTEGRIN ALPHA-11.
FT	DOMAIN	23 1142	INTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1143 1165	POTENTIAL.
FT	DOMAIN	1166 1189	CYTOPLASMIC (POTENTIAL).
FT	REPEAT	38 94	FG-GAP 1.
FT	REPEAT	102 163	FG-GAP 2.
FT	DOMAIN	167 345	WVFA.
FT	REPEAT	359 420	FG-GAP 3.
FT	REPEAT	422 475	FG-GAP 4.
FT	REPEAT	477 537	FG-GAP 5.
FT	REPEAT	539 598	FG-GAP 6.
FT	REPEAT	601 653	FG-GAP 7.
FT	DOMAIN	1154 1162	POLY-LEU.
FT	DOMAIN	1174 1177	POLY-ARG.
FT	CA_BIND	551 559	POTENTIAL.
FT	CA_BIND	613 621	POTENTIAL.
FT	DISULFID	75 83	BY SIMILARITY.
FT	DISULFID	121 139	POTENTIAL.
FT	DISULFID	129 159	POTENTIAL.
FT	DISULFID	659 668	BY SIMILARITY.
FT	DISULFID	674 729	BY SIMILARITY.
FT	DISULFID	781 787	BY SIMILARITY.
FT	DISULFID	881 893	BY SIMILARITY.
FT	CARBOHYD	82 82	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	95 95	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	291 291	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	331 331	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	358 358	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	449 449	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	462 462	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	528 528	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	642 642	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	694 694	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	857 857	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	894 894	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	973 973	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1032 1032	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1040 1040	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	VARIANT	433 433	V -> M.
FT	VARIANT	524 524	/FTid=VAR_009889.
FT	VARIANT	972 972	R -> L.
FT	VARIANT	1003 1003	/FTid=VAR_009890.
FT	VARIANT	1030 1030	L -> P.
FT	VARIANT	1094 1094	/FTid=VAR_009891.
FT	VARIANT	1094 1094	I -> M.
FT	VARIANT	1094 1094	/FTid=VAR_009892.
FT	VARIANT	1094 1094	Missing.
FT	VARIANT	1094 1094	/FTid=VAR_009893.
FT	VARIANT	1094 1094	L -> V.
FT	VARIANT	1094 1094	/FTid=VAR_009894.
SQ	SEQUENCE	1189 AA; 133609 MW; 6030308A4A4CD52 CRC64;	

Query Match 18.6%; Score 1093.5; DB 1; Length 1189;
Best Local Similarity 28.2%; Pred. No. 5.8e-05;
Matches 348; Conservative 215; Mismatches 502; Indels 167; Gaps 47;

QY	1	FNLDTENAMTFQENARG-FQOSVVQ--LOGSR-VVGAPOEIVAAORGSLYQCDYSTGS	56
Db	23	FMDTRKPRVPIGSRRTAFPGYTVQHQHDSGNKMLVWGALETNGYQKTKCPVHGN	82
QY	57	CEPIRL-----QYPVEAVMSGLSLAATYSPOLLACGPTVHOTCSENYVKGELFL	110
Db	83	CTKLNLRVTLNSVSRKDNRLGLSLATPKONSLFACSLWHEHCSSYTTTGMGRV	142
QY	111	GSNLRQOPQKPEALRGCPQSDSDIAFLVDGSGSIIPHDPRRAKEFTSTVMEQ--LKSK	168
Db	143	NSNFRFSKTAP-ALQRC-QTYMDIVLVLDGNSIYP--WVEVQHFLINILKKEFVIGPG	198
QY	169	TLPSLMQYSEBFRIHFTKFPQNNPNRSLIKPTOLLG-RTHATGIRKVVRELEIN	227
Db	199	IQGVVQGGEDVHEFHLDNYSVKDWEASHLEQRGGTETRTAPGLEPARSFAFO--K	256
QY	228	CARNAPKILITGDGKFGDPLGYEDVIPEDAREGVIYVIGV-----GDAPRSEKSR	282
Db	257	GGRGKAKWIVI TDGSHDSP-DLEKVIQOSRDNVTRYAVAVLGVYNNRRGINPETFLN	315
QY	283	BLNTVASKPRDRHVQINNFEALKTIONBLEKIPATEICTGSSSSPEHEMSQEGSAA	342
Db	316	EIKYIASDDDKHFFNVTDAAALXDIVDALGDRIFSLGNTK-NETSFGLEMSQTGSSH	374
QY	343	ITSNGLPILSTVGSYDWAAGVFLYTSKSK-----STFINMTRVDSMDNDAYLGAAAI	397
Db	375	VVEDGVLLGAVGAYDNGAVLKETSGKVIPLRESYLKEPPEELKNHGAYLGYTVTSVVS	434
QY	398	NRV-OSLVGAPRYOHTGLVAMP-RONTGMSNANVGTQIGAYFGASLCSVDVDSNGS	455
Db	435	SROGRVTVAGAPRFNHTGKILFTMHNNRSITTHQAMRGQIGSYGSEITSDIDGCV	494
QY	456	TDVLIGAPHYEQTR-GGQSVGCLPRGQARQCDAVLYGECQCPWGRGAALTVLGD	514
Db	495	TDVLVGAAMYFNEGRGKVVYVEL---RQNRVYNGTLKDSHSYQNAFGSSIASVRD	551
QY	515	VNGKLTDAIAGPGEEDNRGAVLPHGTSGSGISPSHSQRIAGSKLSPLRYGQCSLSG	574
Db	552	LNQDSYNDVVVVGAPLEDNHAGAIYIFHGFRGS-ILKTPKQRIITASELATGLQYFGCS	610
QY	575	QDQLTMDGLVLTGAOCHVLLRSQPLRVKAMFENPREVANVF--ECNDQVVKGE	632
Db	611	QLDNEGLDLAVGALGNVILNSRPVQINASLHFEPSKI-NIFHRDC-----KR	661
QY	633	AGEVRVCL-----HVQSTRDLRREGQIOSVYTDIALDSGRPHSRVAFNET	679
Db	662	SGRDATCLAAFLCPTPIFLAPHFQTTVG-----IRYNATMDERRYTPRAHLDG	711
QY	680	KNS-TRBQTVLGLTQCTCLKLPNCIEDPVSPIVLRNLFSLVGTPLSAGNLRPVL	738
Db	712	GDRFTNRAVLSSQGLCERINFHVL-D-TADYVQVTFVSVEYSLEDP-----DHG	764
QY	739	EDAQRLLTALPFPFKNCNDNICQDDL-----SITFSF	771
Db	765	DGWPPTLRVSVFPWNGCNEDEHCVPLDLVDARSDLPTAMEYCORVLRKPAQCSAYTLSP	824
QY	772	MSLDCLVVGPREFNVTVTVENDGEDSVYTOVTFPPFLDLSYRKVSTLONORSQSRWLA	831
Db	825	DTVFIIESTQRVAVENTNENGENAYSTVLNLSQANLQF--ASLIQKSDSGS--IE	880
QY	832	CESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNLKLKANVTSENN	891
Db	881	CVNEER-----RLQKQVCNVSYPFFRAKAKVAFRLDPEF-SKSIHLHLEIELAAGSDN	934
QY	892	MPTNKTE--POLSLPVKIAYVMVWTSHGVSSTKY---LNFAS--ENTSRVMQHQYQVSN	944
Db	935	ERDSTKEDNVAPLPHLYEADVLFTRSSSLSHYEVKLSLSERYDYGIPGPPFCIFRION	994
QY	945	LQO---RSLPISLVFLVPVRLNQTVWDRPOVTFSENLSSTC---HTKERLPSSHDFLA	997
Db	995	LGLPFIHIGIMKIITPATRSGRNLLKLRLDPLT-DEVANTSCNIGWNSSTERYPVE--E	1051
QY	998	ELRKAPVWNCISIAVCQRIQCDIPFGIQEBEFNATIKGNLSPDMY-----IKTSNHL	1052

Db 1052 DRRAPQLKSHSDVVSINCNRLVP-NOEINFLHGLND---WLSRLKALKYKSKWIMVN 1107
 Qy 1053 STAEILFNDVSFTLLPGQAFVRSQTETKVEFEVFN-----PLPLVGVSGVGLLLA 1106
 Db 1108 AALQRFH-SPF-----IFREEDPSQIVFELSKQEDWQVPTIIVGSLGGLLLA 1158
 Qy 1107 LITAAIYKLGFKF-KROYKDMSEGGPGCA3P 1136
 Db 1159 LVLALWKLGFPRSAARRRE-----PGLDLP 1183

RESULT 11
 ID ITA2 BOVIN STANDARD; PRT; 1170 AA.
 AC P53710;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GP1a)
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b) (Fragment).
 GN ITGA2.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94193647; PubMed=7511592;
 RA Kamata T., Puzon W., Takada Y.;
 RT "Identification of putative ligand binding sites within I domain of
 integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).";
 RL J. Biol. Chem. 269:9659-9663(1994).
 CC !- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
 COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
 RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN
 COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
 CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE
 EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
 EXTRACELLULAR MATRIX.
 CC !- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
 CC !- ASSOCIATES WITH BETA-1.
 CC !- SUBCELLULAR LOCATION: Type I membrane protein.
 CC !- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC !- SIMILARITY: Belongs to the integrin alpha chain family.
 CC !- SIMILARITY: Contains 1 VWFA domain.
 CC !- SIMILARITY: Contains 7 FG-GAP repeats.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L25886; AAB59255.1; -.
 CC PIR; I45914; I45914.
 CC HSP; P17301; LAOX.
 CC InterPro; IPR000413; Integrin_alpha.
 CC InterPro; IPR002035; VWFA.
 CC Pfam; PF01839; FG-GAP; 3.
 CC Pfam; PF00357; integrin_A; 1.
 CC Pfam; PF00092; vwa; 1.
 CC SMART; SM00191; Int_alpha; 5.
 CC SMART; SM00327; VWA; 1.
 CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 CC PROSITE; PS00234; VWFA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Platelet; Signal; Polymorphism; Calcium; Magnesium.
 FT NON_TER 1 1

FT SIGNAL <1 18
 FT CHAIN 19 1170
 FT DOMAIN 19 1121
 FT TRANSMEM 1122 1143
 FT DOMAIN 1144 1170
 FT REPEAT 34 92
 FT REPEAT ? ?
 FT DOMAIN 177 367
 FT REPEAT ? ?
 FT REPEAT 423 475
 FT REPEAT 477 538
 FT REPEAT 540 599
 FT REPEAT 604 656
 FT CA_BIND 488 496
 FT CA_BIND 552 560
 FT CA_BIND 616 624
 FT SITE 472 474
 FT SITE 1146 1150
 FT DISULFID 72 81
 FT DISULFID 669 726
 FT DISULFID 778 784
 FT DISULFID 854 865
 FT DISULFID 1008 1039
 FT DISULFID 1044 1049
 FT CARBOHYD 94 94
 FT CARBOHYD 101 101
 FT CARBOHYD 332 332
 FT CARBOHYD 421 421
 FT CARBOHYD 449 449
 FT CARBOHYD 464 464
 FT CARBOHYD 588 588
 FT CARBOHYD 748 748
 FT CARBOHYD 945 945
 FT CARBOHYD 1063 1063
 FT CARBOHYD 1070 1070
 FT VARIANT 580 580
 FT VARIANT 588 588
 FT VARIANT 725 725
 SQ SEQUENCE 1170 AA; 138929 MW; ECEPFC5P2448FB1 CRC64;

Query Match 18.5%; Score 1085; DB 1; Length 1170;
 Best Local Similarity 27.6%; Pred. No. 2.1e-64;
 Matches 335; Conservative 217; Mismatches 495; Indels 168; Gaps 47;

Qy 1 FNLDTEANMTQ-ENARFGQSVQL-----QSRVVVVGAPQIVAAVNRGSLVQC--DYST 54
 Db 19 YNVGLPKAKIFSGPSSEQGYAVQVQFINPKGNWLLVSGPSGPKNKGMDVTKCPDUST 78
 Qy 55 GSCBPRLQ-----VPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGIC 107
 Db 79 TCEKLNLTSTSMNSNTMTNLSGLTLTENVGTGGFLTCGLPWAQCGSQVYTTGVC 138
 Qy 108 FLFGNLRQOKPEALRQPOEDSDIAFLVDGSGSIIPHDPRAKEFISTVMEQLK-- 165
 Db 139 SDVSPDF-QLRTSPAPVQTCF-SFIDVVVVVCDENSIYPWD--AVKNFLKFPVGLDIG 194
 Qy 166 KSKTLFSLMOYSEFRIFHTFKBFQNNENPRSLIKPITQLL---GRTHATGIRKVVRE 221
 Db 195 PTKQMGILQVANNPRVFNLTFSKSD--EMIKATSQTFQYGGDLNTFKAIQVARDT 251
 Qy 222 LFNITNGARKNAFKILILITDGEKPGDFGVYEDVIEADREGVIRYVIGV-----GDAPR 276
 Db 252 AYSTAAGCRPGATKVMVVVTDGESH-DGSKLKAVIDQCNKILFLFGIAGLYLNALD 310
 Qy 277 SEKSRQELNTVASKPDRDHVFQINNFEALKTIQNLREKIFAIEGTQTGSSSPHEMSQ 336
 Db 311 TKNLKEIKAIASIPTEHFFNVSDADLLEKAGTIGEQIFSIETGVG-GDNFQKMSQ 369
 Qy 337 BGFSAAIT--SNGPILLSTVGSYDNAGGVLYTSKEKSTFINMT--RVDSDMN-DAYLGYA 391
 Db 370 VGFSAEYSPQNNILMILGAVGAYDSGTVQVQKTPHGLIFSKQAFQILQDRNHSYLGVS 429
 Qy 392 AAILLRNVQSLVLGAPRYQHIGLVAMFRQNTGMEVSNANV-----KGTQIGAFYGASL 445

Db 430 VASISTGNSVHFVAGAPRANVTGQIVLYSVN-----ENGNTVTIQSGRDQIGSYFGSVL 484
Qy 446 CSVDVDSNGSTDLVLGAPHYEOTR--GGQVSCVPLPGORARWQCDVAVLYGEOQCPWG 503
Db 485 CAVDVNKDTITDVLGAPYMDLKKKEGRVFLPTKQ-ILNWH--QFLEQNGLENA 541
Qy 504 RFGAALTVLGVNDKLTVAIGAPGEDNRGAVLPHGTSGSISPSHSQRIAGS--KL 561
Db 542 RFGSAIAALSDINMDGNDVIVGSPLENQNSGAVIYNHGEHM--IRLRSYQKILGSDRAP 600
Qy 562 SPRLQYFQSGISGQDITMDGLVLTGAGQHVLLRSQPLVKVKAIMBNPREVARNVF 621
Db 601 SSHLYQFGRSLDGGDLNGDSITDVSVGAQGVQVQLMSQSIADVDSASTPKKI--TLL 658
Qy 622 ECNDQVVKGEAGEVRYCLHVQKSTRDRLEGGIQSVVYDIALD-----SGRPHSRAVFN 677
Db 659 NKNAEI-----KLKLP-----SAKRPYNNQNVAVIYNIDBQFSSRVISRLGFK 707
Qy 678 ETKNSSTRQTVGLTQTCR--TLKQLPNCIEDPSPVILRNFSI--VGTPLSAPGNL 733
Db 708 ENNERCLQKTMVISOQRCSEYIIHQEPS---DIISPLMLCMNISLENPT----- 756
Qy 734 RPLVLAEDAQRLFTALFFPEKXNCGNDNICQDLSITF----SFMSLDCLVVGGRPRENVTV 789
Db 757 NPALAEYSEIVKVFISIPFKDQDGGVICISDLVNLVQQLPATQQCPPIVSNQKRLTFSV 816
Qy 790 TVRNGEDSTRTQVTFPPDLDSYRKVSTLQNRQSRQSRWLACESAST--EVSQALKSTS 848
Db 817 QLRNKKESAVNTEIVDFSENLF-----ASWNPVNDGTETVCQIASQKSVT 864
Qy 849 CSINHPIPPENSEVFNITPDVDSKASLGNKLLKANVTSENNMPTNKTEFOLELPVKY 908
Db 865 CNVGYPAKSKQVTFINFPDNLQ-NLQNASISPRALSESQENMADNSVNLKLSLLY 923
Qy 909 AVYMWVTHGVSTKYLMFTASNTSRVMOHQYVSNLQOR-----SLPISLVPLV 958
Db 924 DAEIHIT-RSTNINIFYSLDGNVSVV-HSFE--DIGPKFIPSIKVTGSGVPVSWA--- 976
Qy 959 PVRLNQTVMDRPOVTFSEN--LSSYCHTKE-----RLPSHSDFLAE- 998
Db 977 -----SVLIHPQYTKDKNPLMVLTVHTDQAGDISCEAEINPLKIQTSSSVSFKSEN 1030
Qy 999 LRKAPVWNCISAVCQRIQCDIPFGIOEENFATLKNLSPDWYIKTSHNHLIVSTAEI- 1057
Db 1031 PRHIKELNCRATSCSNIMCLWRLDQVKGEYFLNSTRWNGTFAASTFQTVQLTAAAEID 1090
Qy 1058 LFNDVSFTL-----LPGQGFVRSQETKVERPE-VPNPLPLVSGSVGLLALLLITA 1110
Db 1091 TYNQIVVIENTVTIP-----LTIMKPKHEVPTGVIVGSVIAGILLALLAIVA 1140
Qy 1111 ALYKLGFFKQYKDM 1125
Db 1141 ILWKLGFFKRYEKM 1155
RESULT 12
ITA2 MOUSE STANDARD; PRT; 1178 AA.
AC Q62469; Q62163;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 43, Last annotation update)
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GPIa)
GN (Collagen receptor) (VLA-2 alpha chain) (CD49b).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=lung;

RX MEDLINE=94363406; PubMed=8081889;
RA Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,
RA Deljanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.;
RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but
RT not virus binding.";
RL Cell Adhes. Commun. 2:131-143(1994).
RN [2]
RP SEQUENCE OF 450-1178 FROM N.A.
RC TISSUE=lung;
RX MEDLINE=94355691; PubMed=7521231;
RA Wu J.E., Santoro S.A.;
RT "Complex patterns of expression suggest extensive roles for the alpha
RT 2 beta 1 integrin in murine development.";
RL Dev. Dyn. 199:292-314(1994).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING
CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
CC COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN
CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION
CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z29987; CAAS2877.1; -;
DR EMBL; X75427; CAAS3178.1; -;
DR PIR; S44142; S44142.
DR HSP; P17301; IAOX.
DR MGD; MGI:96600; Itga2
DR InterPro; IPR004433; Integrin_alpha.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWF_1
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Platelet; Signal; Repeat; Calcium; Magnesium.
FT SIGNAL 1 26
FT CHAIN 27 1178 INTEGRIN ALPHA-2.
FT DOMAIN 27 1129 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1130 1151 POTENTIAL.
FT DOMAIN 1152 1178 CYTOPLASMIC (POTENTIAL).
FT REPEAT 42 100 FG-GAP 1.
FT REPEAT 185 375 FG-GAP 2.
FT DOMAIN 375 375 VWFA.
FT REPEAT 431 483 FG-GAP 3.
FT REPEAT 485 546 FG-GAP 4.
FT REPEAT 548 607 FG-GAP 5.
FT REPEAT 612 664 FG-GAP 6.
FT REPEAT 696 704 FG-GAP 7.
FT CA_BIND 496 504 POTENTIAL.
FT CA_BIND 560 568 POTENTIAL.
FT CA_BIND 624 632 POTENTIAL.
FT SITE 480 482 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 1154 1158 GPFKR MOTIF.
FT DISULFID 80 89 BY SIMILARITY.

FT	DISULFID	677	734	BY SIMILARITY.
FT	DISULFID	786	792	BY SIMILARITY.
FT	DISULFID	862	873	BY SIMILARITY.
FT	DISULFID	1016	1047	BY SIMILARITY.
FT	DISULFID	1052	1057	BY SIMILARITY.
FT	CARBOHYD	102	102	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	109	109	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	429	429	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	457	457	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	472	472	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	696	696	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1054	1054	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1071	1071	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1078	1078	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE	1178	1178	AA; 128926 MW; 1F194B9C0240F465 CRC64;

Query Match		18.2%;	Score 1071;	DB 1;	Length 1178;
Best Local Similarity		27.9%;	Pred. No. 1.8e-63;		
Matches 342;		Conservative 208;	Mismatches 487;	Indels 188;	Gaps 44;
Qy	1	FNLDTENMTFQ-ENARGFQSVVOL---	QGSRVVTCAPQEI	VAANQGSLSYQC--DYST	54
Db	27	YNYGLPKAKIFGSSSQFGYSVQQLTNPQGNLLV	GPSWGPENR	MGVDIVKCPVDLPT	86
Qy	55	GSCEPIRLQ-----VPVEAVNMSGLSLAAATTS	PPQLLACQPTVHQT	CSNTYVKGLC	107
Db	87	ATCEKLNQNSASISNVTEIKTNSLGLTLTRNPGTG	FLTCPLMAHQCNQY	YATGIC	146
Qy	108	FLFGSNLRQOQP---KPEALRCQPQEDSDIAFLV	DGSGSIIIPHD	FRAKEFTSYMEQL	164
Db	147	---SDVSPDFQFLTSPVQACPSL--VDWVVVCD	ESNIYP--WEAVK	NFLVFTVGL	199
Qy	165	K--KSKTLFSLMOVSEEFRIHFTPKBFQNNPNR	SLIKPITQLLG--RTH	TATGIRKVVRE	221
Db	200	DIGPKTKTOVALIQVANEPRIIIFNLNDETRED	MVQATSETRQHGDLTNT	FRALPARDY	259
Qy	222	LPMITNGARKNAFKILILI	TGDKFQDPLGYEDVI	PEADREGVIRYIGV----	276
Db	260	AYSQTSQGRPGATKVMVWVITDGESH--DGSKL	KTVIQCCNDDEILR	FGVLGYLNRAWD	318
Qy	277	SEKSRQELNVTASPKPRDHVFOINNEPALKTIQ	NOLREKJFAIBGTQTGSSS	PEHEMSQ	336
Db	319	TKNLIKETIKALSTPTERYFNVNDAEALALEK	AGTLGEQIFSIESTVQG--	GDNPFQMEWAQ	377
Qy	337	EGFSA--AITSNGPLLTGVSYDWAGGVFYAT	SKEKSTFIMNT--RVDSD	MN-DAYLGYA	391
Db	378	VGFSAVAPQNDILMLGAVGAFDWSGTLVQET	SHKPIFPKQAFQDVLQ	DRNHSSFLGYS	437
Qy	392	AAIILNRVQSLVTCAPRYQHI	GLVAFPRQNTGHWES	NANV----KGTQIGAV	447
Db	438	VAAIISTEDGVHFVAGAPRANTVGOI	VLVSVNK---QGNVT	VIQSHRGDQIGSYFGSVLCS	494
Qy	448	VDVDSNGSTDLVLICAPHYEQTR--GGQVSV	CPLPQGRARWCDAVLYE	QOQPWGRF	505
Db	495	VDVKDITITDVLVAGPYMNDLKKEGKVLYLT	ITTKGILMQHQ---FLEG	PEGTGNARF	551
Qy	506	GAALTVLGVNDGDKLTVAICAPGEEDNR	GAVYLFHGTSGSGS	ISQRLAGSKLSR-	564
Db	552	GSATLAALSDINMGDFNDVIGSPVENENSG	AVIYNGHQGT-IRTKYSOK	LIGSNGAPFR	610
Qy	565	-LQYFGQSLSGQDITMDGLVDLTVGAOCHV	LLILRSQPLRVKAKIMEFN	PREVARNVPEC	623
Db	611	HLQFGRSLDGVGLNGDSITDVSIGALQ	GVILQWSQSIADVAIEAL	FTP-----	660
Qy	624	NDQVYKGEAGEVRVCLHVQKSTRDLREG	QIQSVVTVDLALD----SGRPH	SRVAFNET	679
Db	661	-DKITLLNKADAKITLKCRAEAFR	PAGQNNQV--AILEFMTL	DADGHSRYS	717
Qy	680	KNSTERQTVLGLTOTCET--LKQLPNC	IEDPVSPIVLRILNFS	LVGTPLAGF	737
Db	718	SERFLQKNWNVNEVKCEHHISIQNPS---	DVNPLDLRVDISLEN	PQTS-----PAL	768
Qy	738	AEDAQRLETFALFPFPEKNCNCDNICOD	DLISI-----TFSM	SLDCLVWGGPREFNV	788

alpha 2).";
 RL J. Clin. Invest. 92:2427-2432(1993).
 [5]
 RP VARIANT GLU-534.
 RX MEDLINE=20206009; PubMed=10744142;
 RA Kroll H., Gardemann A., Tschert A., Haberbosch W., Santoso S.;
 RT "The impact of the glycoprotein Ia collagen receptor subunit A1648G
 gene polymorphism on coronary artery disease and acute myocardial
 infarction.";
 RL Thromb. Haemost. 83:392-396(2000).
 CC -I- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
 COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND B-CADHERIN. IT
 RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-G-E-R IN
 COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
 CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE
 EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
 EXTRACELLULAR MATRIX.
 CC -I- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2
 associates with beta-1. Interacts with HPSS.
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -I- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -I- POLYMORPHISM: Position 534 is associated with platelet-specific
 allonantigen HPA-5 (BR). HPA-5A/BR(A) has lys-534 and HPA-5B/BR(B)
 has Glu-534. HPA-5B is involved in neonatal alloimmune
 thrombocytopenia (NAIT or NATP). The K534E polymorphism may play a
 role in coronary artery disease (CAD).
 CC -I- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -I- SIMILARITY: Contains 1 VWFA domain.
 CC -I- DATABASE: NAME=PROW; NOTE=CD guide CD49b entry;
 WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49b.htm".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 DR EMBL; X17033; CAA34894.1; --
 DR EMBL; AF512556; AAM34795.1; --
 DR PIR; A33998; A33998.
 DR PDB; LAOX; 25-NOV-98.
 DR PDB; 1DZ1; 02-AUG-01.
 DR Genew; HGNC:6137; ITGA2.
 DR MIM; 192974; --
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0005886; C:plasma membrane; TAS.
 DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 DR GO; GO:0005518; F:collagen binding; TAS.
 DR GO; GO:0007596; P:blood coagulation; TAS.
 DR GO; GO:0007160; P:cell-matrix adhesion; TAS.
 DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWFA.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR SMART; SM00191; Int alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS50234; VWFA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;
 KW 3D-structure.
 FT SIGNAL 1 29
 FT CHAIN 30 1181 INTEGRIN ALPHA-2.
 FT DOMAIN 30 1132 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1133 1154 POTENTIAL.
 FT DOMAIN 1155 1181 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 1155 1161 INTERACTION WITH HPSS.

FT REPEAT 45 103 FG-GAP 1.
 FT REPEAT ? ? FG-GAP 2.
 FT DOMAIN 188 378 VWFA.
 FT REPEAT 378 433 FG-GAP 3.
 FT REPEAT 434 486 FG-GAP 4.
 FT REPEAT 488 549 FG-GAP 5.
 FT REPEAT 551 610 FG-GAP 6.
 FT REPEAT 615 667 FG-GAP 7.
 FT CA_BIND 499 507 POTENTIAL.
 FT CA_BIND 563 571 POTENTIAL.
 FT CA_BIND 627 635 POTENTIAL.
 FT SITE 1157 1161 GPFKR MOTIF.
 FT DISULFID 83 92 BY SIMILARITY.
 FT DISULFID 680 737 BY SIMILARITY.
 FT DISULFID 789 795 BY SIMILARITY.
 FT DISULFID 865 876 BY SIMILARITY.
 FT DISULFID 1019 1050 BY SIMILARITY.
 FT DISULFID 1055 1060 BY SIMILARITY.
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 343 343 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 699 699 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1057 1057 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1081 1081 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 534 534 K -> E (IN ALLOANTIGEN HPA-5B;
 gbsnf:1801106).
 FT TURN 170 170 /FTIG=VAR_003977.
 FT STRAND 173 180
 FT TURN 183 184
 FT HELIX 188 199
 FT TURN 201 201
 FT STRAND 204 204
 FT TURN 206 207
 FT STRAND 209 216
 FT STRAND 220 224
 FT TURN 226 228
 FT HELIX 232 240
 FT TURN 241 241
 FT HELIX 252 262
 FT TURN 263 264
 FT HELIX 266 268
 FT TURN 269 269
 FT STRAND 275 282
 FT HELIX 289 291
 FT TURN 292 301
 FT TURN 302 303
 FT STRAND 304 311
 FT HELIX 313 317
 FT TURN 318 319
 FT HELIX 323 330
 FT TURN 331 332
 FT HELIX 337 340
 FT STRAND 341 344
 FT HELIX 347 353
 FT HELIX 354 362
 FT TURN 363 363
 SQ SEQUENCE 1181 AA; 129295 MW; 7E1B7ED968A94070 CRC64;
 Query Match 18.2%; Score 1068; DB 1; Length 1181;
 Best Local Similarity 26.9%; Pred No 2.9e-63;
 Matches 329; Conservative 214; Mismatches 494; Indels 188; Gaps 43;
 QY 1 FNLDTENAMTPQ-ENARGFGQSVVQL---QSGRVVVGAPQEIIVAAQRGLVQC---DYST 54
 Db 30 YNWGLPEAKIFSGPSSEQFGYAVQOQFINPKGNWLLVGSFSPENRMDGVKCPVDLST 89
 QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSPQLACGPTVHQTCSENTYVKGLC 107

Db 90 ATCEKLNQTSINPVTEMKTNMGLILTRNMTGGELTGGPLWAQCCNQYTTGVC 149
 QY 108 FLFGNLRQPKQFPEALRGCEQEDSDIAFLVDGSGSIIPHDFFRAKEFISTWBEQK-- 165
 Db 150 SDISPDF-QLSASFPAQPCFSL-IDVWVWCDESIVPMD--AVKNLEKFKVQGLDIG 205
 QY 166 KSKTLFSLMOYSEERHIFPKPEQPNPNRSLIKETLQGL-RHTATGIRKVVRELFN 224
 Db 206 PKTQVGLIYANNPRVFNMLTYTKEMIVATSTQSYGGDLNTPGALQYARKAYS 265
 QY 225 IYNGARKNAFKILITIDGKEGDLPLGVEDYIPEADREGVIRYVIGV-----GDAPRSEK 279
 Db 266 AASGGRRSATKVMVVVTDGESH-DGSMKKAVIDQCNHDMILAFGLAVGLYNALDFKN 324
 QY 280 SQQLNTVASKPRPHVFNQFNFAKLTIONQRLKIPALIEGTQTGSSSSFEHNSQGF 339
 Db 325 LIKEIKATASIFTERFYFNVSDEALLEXAGTIGQIFSIETVQG-GNFMOMENSQGF 383
 QY 340 SAAITSNGP--LLSTVGSVDWAGGVFLYTSKEKSTFINNT--RVDSDMN-DAYLGYAAAI 394
 Db 384 SADYSSQNDILMLGAVGAGWSGTIVQKISHGLIFPKQAFDQILQDRNHSYLGVSVA 443
 QY 395 ILNRVQSVLVGAPRYQHOHGLVAMFRONTGMESNANV-----KGTQIGAVFGSLCSV 448
 Db 444 ISTGSTHFAVAGPRANYTGQIVLVSNN-----ENGNITVIOAHRGDOQIGSVFGLCSV 498
 QY 449 DVDSNGSTDLVIGAPHYYEQTR--GGQVSVCPPLRGQARQCDVAVLGEQ-QPWGRFG 506
 Db 499 DVDKTIITDVLVAGPMTYMSDLKEGRVYLTIKKILGQHQ-----FLEGPEGIENTRFG 555
 QY 507 AALTVLGVNGDKLTDVAIGAFCEEDNRGAVLPHGTSGSGISPSHSQRIAGS--KLSPR 564
 Db 556 SAIAALSDINDGCFNDVIVGSPLENGSGAVIYNGHQT-IRTKYSQKILGSDGAFRSH 614
 QY 565 IQYFGSLGGQDLTHDGLVDITVQGNQHVILLRQPVLRVKAIMEFVREVARNVFCN 624
 Db 615 ILYFGSLDGYDLNGDSITDSVIGAFQVQLWQSIAADVAIEASFTEKI--TLVNNK 672
 QY 625 DQVWKGZGAEVRVCLHVQKSTRDRRECOIQSVVTVDLALD-----SGRPHSRVAFNETK 680
 Db 673 AQII-----LKLCE-----SAKFRTKQNNQVAIVNITLADGFSRVTSEGLFKENN 721
 QY 681 NSTROTQVLGTQTC--BTLKQLPNCIEDPVSPVLVNLNLSLGTPLSAGNLRPVLA 738
 Db 722 ERCLQNMVYNAQSCPEHIIYIQEPS-----DVVNSLDLVDLSLENPGTS-----PALB 772
 QY 739 EDAQLRFTALFPFKKNGDNICODDLSTP-----SEMSLDCLVWGPPREFVNTVVRND 794
 Db 773 AYSETAKVSIIPHKDCGEGGLCISDLVLRQIPAAQOPFIVSNQNKRLTFSVTLKXK 832
 QY 795 GEDSVYRTQVTFPLDLVYKXVSTLQNSQSWRLACESAST-EVSGALKSTSCSINH 853
 Db 833 RESAVNTGIWDPSENLF-----ASFSLPVDGTETVTCQVAASQKSVACDVGY 880
 QY 854 PIPENSEVTNITDQVDSKASIGNKLLKANVTSENMTNKTETFOLELPVKYAVYV 913
 Db 881 PALKRQQTFTTINFDNIQ-NLQNSLSFQALSESQENKADNLVNLKIFLLVDAEI- 938
 QY 914 VTSHGVTSKYLNFTASENTSRVQHOYQVSNLQOR-----SLPISLPLV----- 958
 Db 939 ---HLRSTNINFPYELSSDGNVPSIVHSFEDVGPKFIFSLKVTGSGVPVSMATVHIHPQ 995
 QY 959 -----PVLNQTIVNDRQVTF-SENLSSTCTKTKER 988
 Db 996 YTKRNPLMYLTVGQVTDKAGDISCNADINPLKITGTQ-----SSSVSPKSNFR---HTKE- 1047
 QY 989 LPSSHDFLAELKAPVNVGSIACVQIQCIDIPFGIQEEFNATLKNLSFDWYIKTSNH 1048
 Db 1048 -----LNCRTASCSNVTCLWLDVHKGEYFVNVVTVIWNIGTFASSTFQT 1091
 QY 1049 LLIVSTAEI-LFNDVSFTLLPGQAFVRQCTETKVEPPEVNP-----LP--LIVGSSVG 1100
 Db 1092 VQLTAAAEINTYNPFIYVI-----EDNTVTIPLIMKPDKEAEVPTGTGIGSIIA 1141

QY 1101 GLLLLALITAAALKYKGFKKQYKDM 1125
 Db 1142 GILLALLVALKLGFFKKRYKDM 1166

RESULT 14

ITAG_HUMAN STANDARD; PRT: 1167 AA.
 AC 075578; Q90U28;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-10 precursor.
 GN ITGAL0.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Articular chondrocytes;
 RX MEDLINE=98352078; PubMed=9685391;
 RA Camper L., Hellman U., Lundgren-Aakerlund B.;
 RA "Isolation, cloning, and sequence analysis of the integrin subunit
 RT alpha10, a beta1-associated collagen binding integrin expressed on
 RT chondrocytes.";
 RL J. Biol. Chem. 273:20383-20389(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Endothelial cells, and Heart;
 RX MEDLINE=20169197; PubMed=10702680;
 RA Lehnert K., Ni J., Leung E., Gough S.M., Morris C.M., Liu D.,
 RA Wang S.-X., Langley R., Krissansen G.W.;
 RA "The integrin alpha10 subunit: expression pattern, partial gene
 RT structure, and chromosomal localization.";
 RL Cytogenet. Cell Genet. 87:238-244(1999).
 CC -1- SUBUNIT: INTEGRIN ALPHA-10/BETA-1 IS A RECEPTOR FOR COLLAGEN.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-10
 CC ASSOCIATES WITH BETA-1.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Widely expressed with highest expression in
 CC muscle and heart. Found in articular cartilage.
 CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -1- SIMILARITY: Contains 1 VWFA domain.
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF074015; AAC31952.1; --
 CC EMBL; AF112345; AAF21944.1; --
 CC EMBL; AF172723; AAF61638.1; --
 CC HSSP; P17301; 1A0X.
 CC Genew; HGNC:6135; ITGAL0.
 CC MIM; 604042; --
 CC GO; GO:0008305; C:integrin complex; TAS.
 CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 CC GO; GO:0005518; F:collagen binding; TAS.
 CC GO; GO:0007160; P:cell-matrix adhesion; TAS.
 CC InterPro; IPR000413; Integrin_alpha.
 CC InterPro; IPR002035; VWF A.
 CC Pfam; PF01839; FG-GAP; 3-
 CC Pfam; PF00092; vwa; 1
 CC PRINTS; PR01185; INTEGRINA.
 CC PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00191; Int_alpha: 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.
DR PROSITE; PS00234; VWA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Calcium; Magnesium.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 1167 INTEGRIN ALPHA-10.
FT DOMAIN 23 1122 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1123 1145 POTENTIAL.
FT DOMAIN 1146 1167 CYTOPLASMIC (POTENTIAL).
FT REPEAT 38 97 FG-GAP 1.
FT REPEAT 2 7 FG-GAP 2.
FT DOMAIN 167 350 VWA.
FT REPEAT 365 427 FG-GAP 3.
FT REPEAT 428 482 FG-GAP 4.
FT REPEAT 483 545 FG-GAP 5.
FT REPEAT 546 605 FG-GAP 6.
FT REPEAT 608 660 FG-GAP 7.
FT DOMAIN 1134 1140 POLY-LEU.
FT CA_BIND 494 502 POTENTIAL.
FT CA_BIND 558 566 POTENTIAL.
FT CA_BIND 620 628 POTENTIAL.
FT DISULFID 76 86 BY SIMILARITY.
FT DISULFID 666 675 BY SIMILARITY.
FT DISULFID 681 736 BY SIMILARITY.
FT DISULFID 789 795 BY SIMILARITY.
FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 733 733 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 763 763 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 921 921 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1011 1011 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 844 844 I -> L (IN REF. 2).
FT CONFLICT 909 909 G -> V (IN REF. 2).
FT CONFLICT 926 926 E -> D (IN REF. 2).
SQ SEQUENCE 1167 AA; 127573 MW; AE7D3A1C25C1ABE0 CRC64;

Query Match 18.1%; Score 1064.5; DB 1; Length 1167;
Best local similarity 29.0%; Pred. No. 4.9e-63;
Matches 353; Conservative 201; Mismatches 493; Indels 171; Gaps 44;

QY 1 ENLDTENAMTFQENARG-FQGSVVOLOGSR---VVVGAPOEIVAAORGSLYQC----- 50
DB 23 FNLDEHPLPFGPPPEAFQYGVLVHVGQGRQWLVGAFWDPGSDRRGDVTRCPVGGAH 82
QY 51 -----DYSTG-SCBPRLQVPEAVNMSLGLSLAATSPPOLLAGCGTTHQTCSE 99
DB 83 NAPCAKGHLDYOLGNSSHP-----AVNMHLGMSLLETDGCGFMACAPLWSRACGS 134
QY 100 NTYVKGCLPLFGSLNLRQKPEALRGCPQSDSDIAFLVDGSGSTIIPHDPRAXEFTST 159
DB 135 SVFSSGICARVDASFOQGSAPTAOR-CPTY-MDVIVLDGNSIYP--WSEVQTFLNR 190
QY 160 VMEQL--KXSKTLFLSMQYSEEFRIHPTKEFQNNPNRSLIKPIQLLGR-THATGIR 216
DB 191 LVKGLFIDPEQIQVGLVQYGESPVHWSLGDFTKEEVVRAAKNLSRREGRETKTAQAIM 250
QY 217 KVRELPLNTNGARKVAFKILITDEKEF-GDPLGYEDVIEADREGVIRVIGV-GDA 274
DB 251 VACTGFSQSHGGRPEARLLVVVTDGESHDEELPAALKACEAGR--VTRIGIAVLGHY 308
QY 275 FRSEKS----ROBLNIVASKPPRDHVFQINNFEALKTIQNLREKIFAEGTQTGSSSSF 330
DB 309 LRRQDPSSFLRIRTIASDPDRFFENVTDAAALTDIVDALGDRIFLEGSHAEENSSF 368
QY 331 EHEMSQEGSAITSGNPLLSVGVTDWAGGYFLYTSKEKSTFINWTRVDS-----DMND 385

369 GLEMSQIGFSTHRLKDGILFGWVGAYDWGGSVLMLEGGHRLPPRMALEDEFPFALQNH 428
386 AYLGYA-AAIILNRVQSLVILGAPRYQHIGLVAMFR-ONTQGWESNANVKGTQIGAYFGA 443
429 AYLGSVSSMLRGERRLFLSGAPFRHKGVIAPOLKDKGAVRVAQSLQCEQIGSYFGS 488
444 SECSVDVDSNGSTDLVIGAPHY--EQTREGQYVSVCLPFGQARWQCDVAVLGEQGP 501
489 ELCPUDTDROGTTDVLVAAPWFLGPQNKETGRVTYLV--GQSSLLTQGTLOPEPQD 546
502 WGRFGAALTVLGDVNGDKLTVAIAGPEEDNRGAVLFGHTSGSGISPSHSORIASKL 561
547 -ARTGFAMGALPDNLQDGFADVAVGAPLEDHQGALYLHGTQ--SGVRFPHQAQIAASM 604
562 SPRLQYFQGSLSGGQDLTNDGLVBLTVCAQGHVLLRSQVLRVKAMFEPREVARNPV 621
605 PHALSYFGRSYDGRDLDDGLDLDVAVGAQGAAILLSRPIVHLTPSLVETVQALISVQR 664
622 ECNDQVVKGEAG--EVRVCLHVKSTRDRIRREGIQISVVVYDLDLSDGRSPHSRAVNET 679
665 DCR--RGQEAVALTAALCFQVTSRTFGRWDH--QPYMEFTASLDWETAGAAAFDGS 718
680 --KNSTRQTOVLGUTQCTETLKLQPNCEIDPVSVPLRLNLSVLTGTPSLAFGNLRVL 737
719 CQRLSPRLRLSVG-NVTCEQLHFHVLDT-TSDYLRPVALTVTFALDNTTKFG-----EVL 771
738 AEDAQRLFTALPPEKNCNDNI CODDLSITFSFMSLDC-----LVVGGPREFNAVTV 789
772 NEGSPTSQKLVPSKDCGPDNECVTDLVQ---VNYDIRGRKAPFVVRGRRKVLVST 828
790 TVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNSQRSWRSLACSSASSTVSGALKSTSC 849
829 TLENKENAVNTSLIIFSRNL---HLASLTPQR-ESPIKVECAAPSA-----HARLC 877
850 SINHPPEENSEVTNITFDVDSKASLG---NKLL-----LKANVTSENMPRTNKTBFQ 901
878 SVGRPVFTGAKVTFLLIEFEFSCSLLSQVFGKLTASDSLSRNGTLQENTAQT----- 931
902 LELPVKAVYVMTSHGVSTXYLNFMTASENTSRVQHQY-----QVSNLG 946
932 -----SAYIQEYEPH-----LLFSSSETLHRYEVHPYGLPVGPGPEPKTLRVQNLG 978
947 ---QESLPLSLVFLVP-----VRLNQVINDRPQVTFSENLSSCTCHKERLPESH 994
979 CYVVSGLLISA--LLPAVAGHNTFLSLSQVI-----TNNAACVQMLTEPFGPEPV 1027
995 FLALRKAPVYVNCISIAVCQRIQCDIPFGIOEENFATLKGNLSDFYIKTSHNHLIYST 1054
1028 HPEELOHTNRLNGSNTQCVVRCHLGQAKGTEVSGLLRLVHNEFFRAXFKSLTVVST 1087
1055 AEILPNDVFTLLPQOGAFVRSQTEKVEPEVPNPPLIVGSSVGGLLILALITAAVYK 1114
1088 FELGTEGSLVQLTEAGRSWSESLLEV-VQTRPILISLWILIGSVLGGLLALLVFCLEWK 1146
1115 LGFF-----KROYK 1123
1147 LGFPAHKKIPEEKREEK 1164

RESULT 15
ITAL RAT
ID ITAL RAT STANDARD; PRT; 1180 AA.
AC P18614;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 precursor (laminin and collagen receptor) (VLA-1)
DE (CD49a).
GN ITGAL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;

[illegible]

Db 621 PSGDGKTLKFFQSIHGEMDLNGDGLTDVTIGLGGAALEFWARDVAVVKTMPENKV 680
 QY 617 ARNVFECNDQVVKREAG--EVRVCLHVQ-KSTRDLREGQIQSVVTDALDSCRPHSR 673
 Db 681 NQKKNCR---VEGKETVCINATMCFHVKLSKEDSIYEADLQ-----YRVTLDSLRQISR 733
 QY 674 AVFNET-----KNSTRQTOVLGLTQTCETLKQLPNCI-----EDFVSPIVLRL 718
 Db 734 SFSGTOERKIQENITVRESE-----CIRHSFYMLDKDFQDSVRVTL 776
 QY 719 NFSLVGTPLSAGNLRPVLAEOQLFTALFFPKXCGNDNICODDLISITFSFMSLDCLV 778
 Db 777 DFNLT-DPENG-----PVLDDALPNSVHEHIFPAKDCGKERCISDLTLNVSTTEKSLLI 830
 QY 779 VGGPRE-FNVVTVVRNDGDSVRTQVTFPPFLDLSYRKVSTLQNCORSQSWRLACESASS 837
 Db 831 VESQHDKNVSLTVKNGDSANTRTVQHSFNLFSGIEEIQD-----SCESN-- 880
 QY 838 TEVSGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLL-LKANVTSNNMPTN 896
 Db 881 -----QNTICRVGYPPFLRAGETVTKIIFQFNTSHLSENAIHLATSDEEPLESIN 933
 QY 897 KTEFOLELPKYAV---TWVVTSHGVST-----KVLNFTASENTSRVMOHQYQVSNL 945
 Db 934 DNEVNISIFPKYEVGLQFYSSASEHHISVAANETIPEFINST--EDIGNEINVFTIRKR 991
 QY 946 QORSLP---ISLVF-----LVPRLNOTVIND-----RP-----Q 972
 Db 992 GHFPMPELQLSIFPNLTADGYVLYPIC-----WSSSDNVNCRPSRLEDPFGINSCKK 1045
 QY 973 VTFS-----ENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCRIQCDI--- 1019
 Db 1046 MTISKSEVLKRTIQDCSSTC-----GVATITCSLLPSDLSQVNVSL 1088
 QY 1020 ---PFFGIOEEP---NATLKGMLSPDWYIKTSHNHLIVSTAEILFNDSVFTLLPQQGAF 1073
 Db 1089 LMKPTF-IRAHFSSNLTLRGLK-----SENSSLTLSSN----- 1123
 QY 1074 VRSQETKVEPEVPNPPL--IVGSSVGLLILALITAAKYKLGEPKQYKMMSE 1128
 Db 1124 RKZELAIQISKDGLPGRPLWILLSAFAGLLLLMLLILALMKRIGFKPKPKKKMEK 1180

Search completed: June 7, 2004, 17:12:59
 Job time : 21.9719 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:56:30 ; Search time 39.4889 Seconds
(without alignments)
9084.693 Million cell updates/sec

Title: US-09-902-481B-5

Perfect score: 5876

Sequence: 1 FNLTENANTQENARGFGQ.....FKRQVKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL.25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4381	74.6	1151	11	Q9J130 rattus norv
2	3901.5	66.4	1036	11	Q8CA73 mus musculu
3	3809.5	64.8	920	6	Q28984 sus scrofa
4	3480	59.2	1169	4	Q81VA6
5	3320.5	56.5	1169	11	Q9QXH4
6	3244.5	55.2	1161	11	Q9QYB7
7	1534.5	26.1	1161	11	Q9WTV4
8	1524	25.9	1160	11	Q9R200
9	1409	24.0	1196	13	Q98TF1
10	1359.5	23.1	1086	4	Q96HB1
11	1358.5	23.1	1187	13	Q98TF0
12	1278	21.7	927	6	Q8HZV0
13	1167.5	19.9	1167	11	Q88340
14	1123	19.1	1167	11	Q88341
15	1053	17.9	1171	13	O42094 gallus gall
16	1052.5	17.9	1188	11	Q7CQC3 mus musculu

17	1045	17.8	1038	11	Q8BS01
18	1018.5	17.3	1160	6	Q8MKF4
19	1018	17.3	895	11	Q9WUF8
20	865	14.7	348	4	Q8TES5
21	850	14.5	1332	5	Q9BPQ8
22	808	13.8	205	11	Q63001
23	753.5	12.8	780	13	Q06271
24	738	12.6	823	4	Q8WY18
25	686.5	11.7	823	11	Q8CE84
26	669	11.4	1032	11	Q61989
27	643	10.9	1036	11	Q91YD5
28	642.5	10.9	1033	6	Q9BGU3
29	627.5	10.7	1474	5	Q86GB7
30	623.5	10.6	257	11	Q8C270
31	619.5	10.5	1041	5	Q9UB90
32	614.5	10.5	1041	5	Q76378
33	579.5	9.9	1054	5	Q9U6S1
34	578.5	9.8	1034	13	Q98TF7
35	555.5	9.5	1053	11	Q80TP5
36	550	9.4	1033	13	O42598
37	546	9.3	1036	6	Q7YRP8
38	534	9.1	1016	13	Q91779
39	530	9.0	974	11	Q924W2
40	529	9.0	1073	11	Q8CC06
41	528	9.0	1119	5	Q86G88
42	526	9.0	1047	6	Q9MZB6
43	525.5	8.9	1007	6	Q9GK48
44	522.5	8.9	1132	11	Q80Z18
45	512.5	8.7	1034	6	Q9TUN4

ALIGNMENTS

RESULT 1

ID Q9J130 PRELIMINARY; PRT; 1151 AA.
AC Q9J130;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Integrin beta 2 alpha subunit.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Fathallah D.M. Sr., Zerrila K. Jr.;
RT Cloning of the rat CD11b cDNA sequence.
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP268593; AAF81280.1; -
DR HSSP; P11215; 1BHQ.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; EG-GAP; 3_
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWF_A; 1_
SQ SEQUENCE 1151 AA; 126943 MW; 8F785695DA074CA5 CRC64;

Query Match 74.6%; Score 4381; DB 11; Length 1151;
Best Local Similarity 72.7%; Pred. No. 1.3e-315;
Matches 827; Conservative 150; Mismatches 158; Indels 2; Gaps 2;

361	QY	GVFLYTSKEKSTFIINMTRVDSMDNDAYLGAAATILNRVQSLVLGAPRYQHIGLVAMFR	420
377	Db	GAFLYTSKDKVTFINTRVDSMDNDAYLGASAVILNRNVQSLVLGAPRYQHIGLVAMFR	436
421	QY	QNTGWHSNANKGTQIGAYFGASCLSDVDVDSNGSTDVLVLGAPHYHQTREGGQSVCP	480
437	Db	ERFGTWEPHTSIK	450
481	QY	PRGORARWQCDVLYGECQGPWGRFGAALTVELGDVNEKLTIDVAITGAPGEEDNRGAVYLF	540
451	Db	-----	450
541	QY	HGTSGSIGSPHSQRIAGSKLSPRIQYFGQSLSGGDLTMDGLVDLTVGAQGHVLLRSQ	600
451	Db	-----SQRITGAHSPGLQYFGQSLSGGDLTMDGLMDLAVGAQGHVLLRAQ	498
601	QY	PVLRVKAIEMFEPVARVNFECNDVVKGEKAGEVRVCLHVQKSTRRLREGQIQSVVT	660
499	Db	PVLRLEATWEPFKKVARSVFACQEOVLNKKDAGEVRVCLRVKXNKKDLREGDIOSTVT	558
661	QY	YDLALDSGPHSRVFNFTKNSRTRQTVLGTQTCETLUKLOLPCNIEDPVPPIVLRNLF	720
559	Db	YDLALDPGCSRIRAPFDETKNTRRTQTVFLGMQKCEITLKLILPCVDVDSVSPILRLNY	618
721	QY	SLVGTPLSAFGNLRPVLAEDAQRLFTALPFEKNGCNDNI CODDLSITFSEMSLOCLVVG	780
619	Db	TLVGEPLRSFGNLRPVLAEDAQRFFTAMPFEKNGCNDSI CODDLSITWSANGLOTLVVG	678
781	QY	GPREFNVTVTRNDGDSYRTQVTFEPFLDLSYRKVSTTLQNRQSRWML-ACESASSTE	839
679	Db	GPQDFNMSVTRNDGDSYGTQVTVVYPSGLSYRKDSASQNP LTKKPFVFKPAEBSSESSE	738
840	QY	VSQALSKTSCSINHPIPPENSEVFTNITFDVDSKASLGNKLLKANKVTSNNMPTNKTE	899
739	Db	GHGALKSTTNWNIHPIPPANSEVFTNITFDVDSHSGFNKLLKALVASENNMSRTHTKT	798
900	QY	FOLELPVKAYVMVTSKGVSTKYLNFTASENSTRVMQHQYQVSNLQGRSLPISLVFLVP	959
799	Db	FOLELPVKYAIYMLVTSDESSIRYLNFTASEMTSKVIHQYQPNLQGRSLPVSVVFWIP	858
960	QY	VRLANQTVIMDRPOVTSSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIANCORICDI	1019
859	Db	VQINNVTVMHDHPQVIFPSQNLSSACHTEQKSPPSNFRDQLERTPVLNCSVAVCKRIQCCL	918
1020	QY	PEFGIQSEFNATLKGNLSPDWYIKTSHNHLIYSTABILFNDSVFTLLPQCAFVSQTE	1079
919	Db	PSFNTQIFNVTLKGNLSFDWYIKTSHGHLIYVSTELLFNDSAFALLPQESYVRSKTE	978
1080	QY	TKVEPEFVPNPLPLVSGSVGGILLALITAAIYKLGFFKQVKDMSEGGPGAPQ	1137
979	Db	TKVEPYEVHNPVPLVIGSSIGGLVLLALITAGIYKLGFFKQVKDMNEAAPQADAPPQ	1036
RESULT 3			
Q28984	ID	Q28984	PRELIMINARY; PRT; 920 AA.
AC	Q28984	AC	Q28984;
DT	01-NOV-1996	(TrEMBLrel. 01, Created)	
DT	01-FEB-1997	(TrEMBLrel. 02, Last sequence update)	
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)	
DE	CD11b	(Fragment1).	
GN	CD11b.		
OS	Sus scrofa (Pig).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
OX	NCBI_TaxID=9823;		
EN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Lee J.-K., Schook L.B., Rutherford M.S.;		
FL	Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; U40072; AAB16869.1; -.		
DR	HSSP; P11215; 1HQH.		

DR	GO:0003035; C:integrin complex; IEA.	
DR	GO:0004895; F:cell adhesion receptor activity; IEA.	
GO	GO:0007160; P:cell-matrix adhesion; IEA.	
DR	InterPro; IPR000413; Integrin_alpha.	
DR	InterPro; IPR002035; VWF A.	
DR	Pfam; PF01839; EG-GAP; 3.	
DR	Pfam; PF00092; vwa; 1.	
DR	PRINTS; PR01185; INTEGRINA.	
DR	PRINTS; PR00453; VWFADOMAIN.	
DR	SMART; SM00191; Int_alpha; 4.	
DR	SMART; SM00327; VWA; 1.	
DR	SMART; PS02234; VWFa; 1.	
FT	NON_TER	1
FT	NON_TER	920
FT	SEQUENCE	920 AA; 102440 MW; E96CC51E35DD5AC CRC64;
Query Match		
Best Local Similarity 64.8%; Score 3809.5; DB 6; Length 920;		
Matches 727; Conservative 84; Mismatches 103; Indels 1; Gaps 1;		
Qy	118	POKPEALRCPCQEDSDIAFLVPGSGSIIPHDFFRAKEFIETVMEOLKSKTILFSLMOYS 177
Db	1	POKPEALRCPCQEDSDIAFLVPGSGSIIPHDFFRAKEFIETVMEOLKSKTILFSLMOYS 60
Qy	178	EEFRIHPTFKFQNNPNRSLKPIITQLLGRTHATGIRKVKVZELFNITWGARKNAEKIL 237
Db	61	EDFVTHFTFNDKRNPSKLLVIRPIQLLGRTHATGIRKVKVZELFNITWGARKNAEKIL 120
Qy	238	ILATDGEKFGDPLGYEDVIEADREGVIRVYVIGVDAFRSEKSRQELNITVASKPPDRHVF 297
Db	121	VWITDGEKFGDPLGYEDVIEADREGVIRVYVIGVDAFNWKSREELNITASKPCGDHVF 180
Qy	298	QINNFEALKTIQNLREKIFAIETGTGSSSSPEHEMSQSGFAAITNSNGPLLSITVGSYD 357
Db	181	QVNNFEAVKTIQNLQEKTTAIEGTQTGSTSSPECEMSQSGFAAITNSNGPLLGAVGSD 240
Qy	358	WAGGVFLYTSKEKSTFINMTRVDSMDNDVILGYAAAIIILNRVQSLVILGAPRYOHIGLVA 417
Db	241	WAGGAFLMFKPRDVFINTRVDSMDNDVILGAYAVILNRQAQSLVILGAPRYOHTGLV 300
Qy	418	MFRQNTGMWESNANVKGTQIGAYFGASLCSVDVDSNSGTDLVILGAPHYEQTREGGVSV 477
Db	301	MEKQNSGAWEKVADIKSGQIGSYFGASLCSVDVNRDGSDDLILGAPHYEQTREGGVSV 360
Qy	478	CPLPRGQARWQCDVILYGEQGPWGFCAALTVILGEVNGDKLTDVAIGAPGEDNRGAV 537
Db	361	CPLPQG-RAKWQCRVILYGEQGPWGFCAALTVILGEVNGDKLTDVAIGAPGEDNRGAV 419
Qy	538	YLFHGTSGSGISPSHSORLACSLPRLOFGQSLSGQDLTMDGLVLTVGAQGHVLL 597
Db	420	YLFHGTSELGLISPSHSQRIACSLSPRLQYFGQSLSGQDLTMDGLMDLAVGAQGHVLL 479
Qy	598	RSQPVLRYKAIEMFNPREVAENVPECDQVWVGKEAGEVRVCLHVQKSTRDRILREGQIQS 657
Db	480	RSQPVLRYEASWVFKPREVAENVYECQAAKTQIAGEVQVCLQVRKSTWDRLEGGTQS 539
Qy	658	VVTVDLALDSCRPHSAVRFNETHKSTRQTOVLGLTOTCEYLKLQLPNCIEDPVPVILR 717
Db	540	IIITVDLALDGRPHPRVAFBEKTRNRQUTQILGSRKCEHALMLPDCVEDSVTPVILR 599
Qy	718	LNFSVLGTFPLSAFGLNRLPVLAEDAQRFTALFPPEKXNCNDNI CODDLISITFSMSLDCL 777
Db	600	LNFSVLGVPASSFENLRPVLAEDAQRFTALFPPEKXNCNDSI CODDLISITFSMSLDL 659
Qy	778	VVGGPREFNVTVTVNRDGEDSVKQVTEFFPLDLISYRKVSTLQNSORSRWLACESASS 837
Db	660	VVGGPRFLIKVTLTVNRQGEDSVKQVTEFFPDSLSYRKVSTQNSORSRWLACESDVS 719
Qy	838	TEVSGALKSTSCSINHPIFFENSEVFTNITPDVDSKASLGNKLLKANVTSENNPRNK 897
Db	720	TEESTALKSTSCSINHPIFFDENSEVFTNVTVDVNDPAFLGYKLLKANVTSENNPPSNK 779
Qy	898	TEFQLELPPKAVYVAVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQORSILPISLVFL 957

```
Db 780 TEFQLELPVYAVVVVTSLEVSTKYFNFTASEKTRHVIHQYQFNILGQRLKLPISVVWF 839
Qy 958 VVRLNCTVWDRPOVTPSENLSSTCHTKERLPSSDPLAEURKAPVNVNCSIAVCORIQ 1017
Db 840 VVRLNCTVWDRPOVTPSENLSSTCHTKERLPSSDPLAEURKAPVNVNCSIAVCORIQ 899
Qy 1018 DIPFGIOEFNATLKNLSP 1038
Db 900 DIPFGIOEFNATLKNLSP 920

RESULT 4
Q8IVAG PRELIMINARY; PRT; 1169 AA.
AC Q8IVAG6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Integrin, alpha X (Antigen CD11C (p150), alpha polypeptide).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Scrausberg K.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC038237; AAH38237.1; -.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; VWFADOMAIN.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; vwa; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
SQ SEQUENCE 1169 AA; 128521 MW; 17B484FEFC79EB6 CRC64;
```

```
Query Match 59.2%; Score 3480; DB 4; Length 1169;
Best Local Similarity 61.2%; Pred. No. 1e-248;
Matches 691, Conservative 141; Mismatches 291; Indels 6; Gaps 4;
```

```
Qy 1 FNLDTENAMTFQENARFGQSVVQLQSGSVVVGAPQEIIVANQKSLYOCYSTGSCBPI 60
Db 20 FNLDTEELTAFRVDSAGFGDSVVQVANSVVVVGAPQKITAANQGLYQCGYSTGACBPI 79
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHOTCSNTYVVKGLCFGLFSGNLQPOK 120
Db 80 GLQVPPEAVNMSLGLSLAATTSPPQLACGPTVHOTCSNTYVVKGLCFGLFSGNLQPOK 137
Qy 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDPRAKFPISTWBLKSKSLFSLMQYSEEP 180
Db 138 LPVSRQECPRQODIVFLIDGSGSISSRNFAFMVRAVISOQRPSPQPSLMQFSNKP 197
Qy 181 RIHPTKFPQNNPNSRLKPIITGLGRHTATGIRKVVRELPTNGARKNAFKILILI 240
Db 198 QHFTFEEPRSRNSPLSLASVHQLQGFYTTATLQNVVHRLPHASYGARRDAKILIVI 257
Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNATVASKPPDRVFN 300
Db 258 TDGKRGDSLDYKDVIPWADAAGIIRYAIGVGLAFQNRNSWELNDIAKSPQEHFKVE 317
Qy 301 NFEALKTIONLREKIPATECTQTSSTSSPEHMSQEGFSAITNSGPISTVGSYDNAG 360
Db 318 DFDALQKIQNLKELFAIEGTETTTSSSFELEMAQEGFSAVFTPDGPIVAGVSTWSG 377
```

```
Qy 361 GVFLYTSKEKSTFTNMTTRVDSMDNDAYLGAAAIILNRRVQSLVLAGPRYOHIGLVAMFR 420
Db 378 GAFLYPNMPTFTINMQENVDMDSDYLGSTELALAKGVQSLVLAGPRYOHIGLVAMFR 437
Qy 421 QNTGWESNANVKCTOIGAYFGASLCSVDVDSNGSTDLVILIGAPHYEQTRGGQSVVCP 480
Db 438 QVSRQWKKAEVTQTQISYFGASLCSVDVDSNGSTDLVILIGAPHYEQTRGGQSVVCP 497
Qy 481 PRQARWQCDVILYGBQGPWGFGAALTIVLGDVNGDKLTDVAIGAPGBEDNAGVYLP 540
Db 498 PRGWR-RWMCDAVLYGQGHGPFGRFGAALTIVLGDVNGDKLTDVVGAPGKENGAYLP 556
Qy 541 HGTSGSGISPSHSORIASKLSPLOYGQSLGGQDLTMDGLVDLTVAAGHVLRLRSQ 600
Db 557 HGVLPSPISPSHSORIASKLSPLOYGQSLGGQDLTMDGLVDLTVAAGHVLRLRSQ 616
Qy 601 PVLRYKAIMENPREVARNYFECNDQVVKGEAGEVRVCLHVQKSTRDLREGEQISVVT 660
Db 617 PVLWVGVMQPIPAEIPSAFECEQVVBQTLVQSNICLYIDKRSKNLLSGRDLQSSVT 676
Qy 661 YDLALDGRPHSRVAFNETKSTRQVNLGTCTETLKLQPNCLIEDPVSPTVLRLNF 720
Db 677 LDALDLPRLSPRATFOETKRSLSRVRLGLKACENFNLLPSCVEDSVTPITLRLNF 736
Qy 721 SLVGTPLSAFGNLRPVLAEADAQRALTALFPFKKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 TLVCKPLAFENLRPMLAADAQRYFTASLPFKKNCADHICQDNLGIFSFPGLKSLVVG 796
Qy 781 GPRENTVTVVNDGEDSYRQVTFPPFLDLSYKVTQLQNRQSRWRLACESASSTEV 840
Db 797 SNLEINAEVWVNDGEDSYGTTVTFSPAGLSYVVAEGQKQGLRSILHLTCDSPAVG-- 854
Qy 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKMLLKANVTSENNMPTNKTPE 900
Db 855 SQGTWSTSCRINHPIFRGGAQITPLATFDVSPKAVLGDRLLLTANVSNNTPRTSKTF 914
Qy 901 QLELPVYAVVTVSHGVSTKYINLTAS-ENTSRVWQHQQVQVNLGQSLPISLVLVFP 959
Db 915 QLELPVYAVVTVVSSHQEQTKYLNFSSESEKSHVAMHRYQVNLGQSLPISVINFVP 974
Qy 960 VRLNQTVIDRPOVTPSENLSSTCHTKERLPSSDPLAEURKAPVNVNCSIAVCORIQ 1019
Db 975 VELNQEAVMQDVEVSHFPQNSLRCSSEKIPASDFLAHTQKPNVDCSIAAGCLFRCDV 1034
Qy 1020 PFGIOEFNATLKNLSPDWYIKTSNNHLLIVSTABILNDSVFTLLPGCAFVRSOTE 1079
Db 1035 PSFVQSELDFTLKNLSFGWVRQILQKVSVSVVSAVITDTSVYSQLEPGQEAFFRAQT 1094
Qy 1080 TKVEPFFVPLPLIVGSSVGGILLLALITAAALYKLGFKRQYKDMWE 1128
Db 1095 TVLEKYKHNPPTPLIVGSSIGGLLLALITAVLYKVGFFKRYKEMME 1143
```

RESULT 5

```
Q9QXH4 PRELIMINARY; PRT; 1169 AA.
AC Q9QXH4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Leukocyte adhesion glycoprotein p150,95 alpha integrin subunit.
GN ITGA5.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,
RA Teuchiya H.;
RT "Isolation of Genes Selectively Expressed by Dendritic Cells.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBSJ databases.
EMBL; AF211864; AAF23492.1; -.
```



```
Db 21 NLDVEPIVFPREDAASFGQTVVQFGSRLVVGAPLEAVAVNQTGRDLYDCAPATGMCQPIV 80
QY 62 LOVPVAVNMISLGLSAAATSPQALLACGPTVHOTCSENTYVKGCLFLPGSNLRQOPKP 121
Db 81 LRSPLAVNMISLGLSVAITNAQLACGTAQACVKNYAKGSCLLIGSSL-QPIQAV 139
QY 122 PEALRGCPQSDIAFLVDSGSIIPHDPRRAKEFTSTWQELKSKTILFSLMOYSEBP 181
Db 140 PASMEPCPROEMDIAFLIDGSGINQDFAQMDKFVKALMGEPASTSTFLSLMOYSNLIK 199
QY 182 IHTEKFKONNPRSLIPITOLLGRTHATGIRKVVRELEFNITNGARKNAFKILILIT 241
Db 200 THFTIPEFNILDPQSLWDIVOLGLTATGIRTVMBELFHSGKNGSRSAKILILVIT 259
QY 242 DGEKFGDPLGYEDVIEADREGVIRYVIGVGDAPRSEKSRQELNIVASKPPRDHFQINN 301
Db 260 DGOKYRDPLEYSVDVPAADKAGIIRVAIGVDGAFQBPATKELNITGSAAPPQDHVFKVGN 319
QY 302 FEALKTIQOLREKIFAIETGTOSSSSFEHMSQEGFSAATNSGPIILSTVGSYDMWGS 361
Db 320 FAALRSIQOLQEKIFAIETGTOSSSSFEHMSQEGFSAATNSGPIILSTVGSYDMWGS 379
QY 362 VFYLTSEKSTFTINMTRVDSMDMDAYLGAAAAIILNRRVQSLVGLGAPRYQHIGLVAMFQ 421
Db 380 AFLYPENTRPTFNMSQENVMDRDSYLGYSTAVAFKVGVSILILGAPRQHTGKVIPTQ 439
QY 422 NTGWNESNANVKTQIGAFPGASLCSVDVDSNGSTDLVILGAPHYVETRGQGVSCVPLP 481
Db 440 EAEHWPKEVGTQIGSVFASLCSVDVDRDSTDLVILGAPHYVETRGQGVSCVFP 499
QY 482 RQORARWQCDVILYCGQGPWGFAGALTIVLGVNDGKLTDAVAGAPGEDNRGAVYLP 541
Db 500 -GVYRGWQCEATLEGQHGQWGFAGALTIVLGVNDGKLTDAVAGAPGEDNRGAVYLP 558
QY 542 GTSGSGISPSHSQRIAGSLSPLOYFGOSLSGGQDITMDGLVDLVGAGQGHVLLRSP 601
Db 559 GASRLAIMPSFQSVRTGSLSLQYFGOSLSGGQDITMDGLVDLVGAGQGHVLLRSP 618
QY 602 VLKVKALMEFNPREVARNVFECDNDVVVKGKEGEVVRVCLHVOKSTRDLRGIOQSVVTV 661
Db 619 LLKVELSIRFAPNEVAXVQWERTPTVLEAGEATVCLTVHKGSPDLL--GNVQSVRY 676
QY 662 DLALDSGRPHSRVAFNETKNSRTOVLGTOTCETLKLQLPNCIEDPVSPVILRLNFS 721
Db 677 DLALDPGLRLISRAIFDETNCITLGRKTLGLGPHCETVKLLPDCVEDAVSPIILRLNFS 736
QY 722 LVCTPLSAGNLEPVLAEADQRLFTALFPFEKNCNDNICODDLSITPSFMSLDCILVVG 781
Db 737 LVDRSASP--RNLPVLAWSQDHITASLPFEKNCQELCEGDLGFSNFGSLQVLVVG 795
QY 782 PREFNVTYVRNDGDSYRTQVTFPPFLDLSYKVSQTLQNRSSQSRWRLACASSTEYS 841
Db 796 SPBLTVTVTWNEGEDSYGLVAFYYPAGLSYERVTGTQ-QPHQYPLRLACAEPAQBD 854
QY 842 GALKSTCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTETPQ 901
Db 855 --LESSSCSINHPIFREGATFMTITFDVSYKAFGLDRLLRAKASNNKPDNTKTAQ 912
QY 902 LELPVYAVVMVTVSHGVSTKYLNFATSENTR--VMQHOVQVSNLQORSLPLSLVFLV 960
Db 913 LELPVYVTVTLISROEDSTNHNVSFSSGEGREQAHAHYRVNNSPLKLVAVNFWV 972
QY 961 RLNQTVWRPQVTPPENLSS--TCHTKERLPSHSDFLAEKRAKAPVNCIAVCQRIQCDI 1019
Db 973 LLNGVAVWD---VTLSSPAQGVSCVSKMPQNPDPFLTQIRRSVLDSCSIADCLHFRCDI 1029
QY 1020 PFGIOEEFNATLKNLSFDWYIKTSHNLLIIVSTAEILPNDSVFTLLPQCGAFVRSQTE 1079
Db 1030 PSLDIQDELDFILRGNLSPGWSQTLQEKVLLVSEAEITFDTSVYSLQEGEAFLAQVE 1089
QY 1080 TKYSPFEPNPLPLIVGSSVGGILLALITALYKLGFPKQYKDM 1126
Db 1090 TTLEEVVYVEPIFLVAGSSVGGILLALITVVLVYKLGFPKQYKEML 1136
```

RESULT 7

```
Q9WTV4 PRELIMINARY, PRT, 1161 AA.
ID Q9WTV4
AC Q9WTV4;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DE 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Integrin alpha L.
GN ITGAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DZ/2J; TISSUE=Spleen;
RA Ma R.Z., Teuscher C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065902; AAD25885.1; -.
DR HSSP; P20701; ILFA.
DR MGD; MGI:96606; Itgal.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VMP_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VMPADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
SQ SEQUENCE 1161 AA; 128240 MW; 86B102F7B209E431 CRC64;

Query Match 26.1%; Score 1534.5; DB 11; Length 1161;
Best Local Similarity 34.0%; Pred. No. 2.3e-104;
Matches 399; Conservative 217; Mismatches 456; Indels 101; Gaps 37;

QY 1 FNLDTENATQENR-RGFGSVVQLQSRVVGAPQEIIVAAORGSLYQCDYSTGSCRP 59
Db 24 YNLDTRPTQSLAQAGRHFGYQVLQIEDG--VVGAPGE---GDNTGGLYHCRSTSEFCQP 79
QY 60 IRLQVPVAVNMISLGLSAAATSPQALLACGPTVHOTCSENTYVKGCLFLPGSNLRQOPQ 119
Db 80 VSLH--GSNHTSKYLGMLTATDAKGSLLACDPLGSRTCQNTYLSGLCYLFPQSLGPM 138
QY 120 KFPALRGCPQSDIAFLVDSGSIIPHDPRRAKEFTSTWQELKSKTILFSLMOYSEB 179
Db 139 QNRPAVQECMKGVLDVFLFDGSQLDRKDFEKMDFMKDVMRKLNTSYQFAAVQFSTD 198
QY 180 FRIHTFKEP-QNNPNRSLIKPTOLLGRTHATGIRKVVRELEFNITNGARKNAFKILI 238
Db 199 CRTEFTFLDYKQNKPDVLLGSVQPMFLNTTRAINVYVAHVFKESGARPDATKVLV 258
QY 239 LITDGEKFGDPLGYEDVIEADREG-----VIRYVIGVGDAPRSEKSRQELNIVASKP 291
Db 259 IITDG-----EASDKGNISAAHDITRYIIIGIKHFVSQKQTLHFASEP 304
QY 292 PRDHVQIINNFEAKTIONQREKIFAIETGTOSSSSFEHMSQEGFSAATNSGPIILS 351
Db 305 VEEFKILDTPEKLDLTDLQRIYAIETGTRQDLTSFNMLSSSGISADLSKGHVYV 364
QY 362 TVGSDVWAGGVF-LYTSKEKSTFTINMTRVDSMDMDAYLGAAAA-IILNRRVQSLVGLGAPR 409
Db 365 AVGAKNAGGLDLREEDLOGATFVQBPDLTSDVRGGYLYTVANWTSSESSPLLAAGAPR 424
QY 410 YOHIGLVAMFR--QNTGWNESNANVKTQIGAFPGASLCSVDVDSNGSTDLVILGAPHY 467
Db 410 YOHIGLVAMFR--QNTGWNESNANVKTQIGAFPGASLCSVDVDSNGSTDLVILGAPHY 467
```

```

Db 425 YOHVQVLLFOAPEAGGRWNTQKIEGTQIGSYFGGELCSVDLDQDGBAELLIGAPLFF 484
Qy 468 EOTRGQVSVCPPLPGQARWOCDAVLVYCEQOPWGRFGAALTVLGVDVNGDKLTVAICA 527
Db 485 GEORGRVFTY---QRRQSLFEMVSELQDGPYPLGRFGAALTALTDINGDLTDVAVGA 541
Qy 528 PGEEDNRGAVLYFHTGTSGSGISPSHSORIASGLSPRLQYFGQSGGQDLTMDGLVLT 587
Db 542 PLEE--QGAVYIFNGKPG--GLSPQSPQRIQGAQVFPGLRWFGRSIHGVKDLGGDLADVV 598
Qy 588 VQAQGHVLLRSQPVLRVKAIMEFNPVARNVFECDQVVKKEAG--EVAVCLHVQKST 646
Db 599 VPGEGRVVVLSRPPVDVVTLSFSPBI PVHEVCSYSAREEQKHGKVLKACFAIKPLT 658
Qy 647 RDLREGIQSVVYDLDALDSGRPHSRVAFVNETKNSFRQTVGLTCTCTLKLQLPNC 706
Db 659 PQ--FQGLLANLSYTLQDGHMRSGRLPFGDSHELSGNTSITP--DKSCLDFHFPFIC 715
Qy 707 TEDVPSPVLRNFSV---GTPLSAFGN--LRPVLAEADAQLFTALFPFKNCGNDNICQ 762
Db 716 IQDLISPINVSNLFSLEEGTPRDQKGRAMQPIILRPSIHTV--TKEI PFKNCGEDKKCE 774
Qy 763 DDLSTTFPMSLDCLVVGSP-----REFNVTVTVRNDGEDSYRTQVTFPPFLDLSYKY 816
Db 775 ANLTSSPARS-----GPLRLMSSASLAVETLNSGEDAYVVRKLDLDPGRGLSPK 827
Qy 817 STLQORSORSRLACESASSTEVSCAL--KSTSCINHPHIFPENSEVTENITFDVDSKAS 875
Db 828 EMLQ---PHSRMPVSCIEL--TEGSSLLTKTLKCNVSSPIFKAGQEVSLQVMENTLLNS 882
Qy 876 LGNKLLKANVTSEN--NMPTNKTEPQLELPVKYAVYVMTVSHGYSKYTLNFTASEN 934
Db 883 WEDFVELNGTVHCENENSSLOEDNSAATHIPVLPVNIUTKEQENSTLYISFTPKGPKTQ 942
Qy 935 VMOHQVQV---SNLQORSIPISLPIVPLVRLNQTIVWDRPQ--VTFSENLSS-----TCHT 985
Db 943 QVQHVYQVRIQPSAYDHNMP--TLEALVGP-----WPHSEDPITYTWSVQDPLVTC 995
Qy 986 KB--RLPSSHDFLAEFLKAPVNVCSIAVCORIOCIPPGIQEAFNATLKNLSFDWYIKT 1044
Db 996 EOLKRESSE--AEQCLPGV-----QFRCPVFF---REILIQVTGTVELSKELXA 1041
Qy 1045 SRNHLIVSTAILFNDSTVFTLLPGQAFVRQSTETKVBFPFVPLPLIVGSSVGGLL 1104
Db 1042 S--STLSLCSLSVSNSSKHFHYLSKA--SEAQVLVKKVLDLIEKEMLHVYVLSGIGGLVL 1099
Qy 1105 LALITAAVLKLGPKPKQYKDMV--SEGPGPQASP 1136
Db 1100 LELIFLALYKVGFFKRNLRKEMADGGVNGSP 1132

```

RESULT 8

```

Q9R200
ID Q9R200 PRELIMINARY; PRT; 1160 AA.
AC Q9R200;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DE 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE Integrin alpha L.
GN ITGA1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spleen;
RA Ma R.Z., Teuscher C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065901; AAD25884.1;
DR HSSP; P20701; 11FA.
DR MGD; MGI:96606; Itgal.
DR GO; GO:0008305; C.integrin complex; IEA.

```

```

GO; GO:0004895; F:cell adhesion receptor activity; IEA.
GO; GO:0007160; P:cell-matrix adhesion; IEA.
InterPro; IPR000413; Integrin_alpha.
InterPro; IPR002035; VWF_A.
Pfam; PF01839; FG-GAP; 3.
Pfam; PF00357; Integrin_A; 1.
Pfam; PF00092; Vwa; 1.
PRINTS; PR01185; INTEGRINA.
PRINTS; PR00453; VMPADOMAIN.
SMART; SM00191; Int_alpha; 5.
SMART; SM00327; VWA; 1.
PROSITE; PS00242; INTEGRIN_ALPHA; 1.
PROSITE; PS50234; VWEA; 1.
SEQUENCE 1160 AA; 128127 MW; A33C531B139F1PAD CRC64;

```

Query Match 25.9%; Score 1524; DB 11; Length 1160;

Best Local Similarity 33.9%; Pred. No. 1.4e-103;

Matches 397; Conservative 216; Mismatches 459; Indels 100; Gaps 36;

Qy 1 FNLDTEANMFOENA--RGFGSVVQLQGRVVGAPQIIVAANQGRSLYQCDYSTGSCPEP 59

Db 24 YNLDTRTQSLAQAGRHFGYQVLIQIEDG--VVGAPGE---GDNTGGLYHCTSEFQCP 79

Qy 60 IRLQVPVEAVNMISGLSLAATTSPPOLLACGPTVHQTCSENTYVKGCLFLFGSNLRQOPQ 119

Db 80 VSLH--GSHNHSKYLGMTLATDAAGSLACDGLSRTCDQNTYLSGLCYLFPQSLGPM 138

Qy 120 KPFEALRGCPQEDSDIAFLVDGSGSIIPHDERRAKBPISTVMBOLKSKTFLSLMOYSEE 179

Db 139 QNRFAVQECMKGVLDLPFLPDGOSLDRKDEKILEFVKDVRKLSNTSYQFAVQFSTD 198

Qy 180 PRIHFTKEF--QNNPNRSLIKPTQLLGRTHATGIRKVVRELFNITNGARKNAFKILI 238

Db 199 CRTEFTFLDYKQNKNDVLLGVSQPMFLINTFRAINVYVAHVKEBSGARPDATKVLV 258

Qy 239 LITGCKFGDPLGTVEDVPEADRG-----VTRYVIGVGDAPFSEKSRQELNATVASKP 291

Db 259 IITDG-----EASDKGNTSAADHITRYIIGIKHFPVQKQTLHTFASBP 304

Qy 292 PRDHFQINNEALNTIQNLREKIFAIEGTQTSSTSSFEHEMSQEGFSAATISNGLPLS 351

Db 305 VEEFKILDTKELKDLPTDLQRRIVALEGTRNRODLTSFNMELSSSGISADLSKGVVVG 364

Qy 352 TVGSYDAGGVF--LYTSKEKSTFNMTKRVDSMDNDAYLGAAA--ILLNRVQSLVGLGAPR 409

Db 365 AVGAKDWAGGDLREDIQGATFVGQBELTSDVRGGYLGTYVAMWTSRSSRPLAAGAPR 424

Qy 410 YOHIGLVAMER--QNTGVMWESNANVKGTQIGAYFGASLCVVDVDSNGSTDLVLGAPHY 467

Db 425 YOHVQVLLFOAPEAGGRWNTQKIEGTQIGSYFGGELCSVDLDQDGBAELLIGAPLFF 484

Qy 468 EQTRGGQVSVCPPLPGQARWOCDAVLVYCEQOPWGRFGAALTVLGVDVNGDKLTVAICA 527

Db 485 GEORGRVFTY---QRRQSLFEMVSELQDGPYPLGRFGAALTALTDINGDLTDVAVGA 541

Qy 528 PGEEDNRGAVLYFHTGTSGSGISPSHSORIASGLSPRLQYFGQSGGQDLTMDGLVLT 587

Db 542 PLEE--QGAVYIFNGKPG--GLSPQSPQRIQGAQVFPGLRWFGRSIHGVKDLGGDLADVV 598

Qy 588 VQAQGHVLLRSQPVLRVKAIMEFNPVARNVFECDQVVKKEAG--EVAVCLHVQKST 646

Db 599 VPGEGRVVVLSRPPVDVVTLSFSPBI PVHEVCSYSAREEQKHGKVLKACFAIKPLT 658

Qy 647 RDLREGIQSVVYDLDALDSGRPHSRVAFVNETKNSFRQTVGLTCTCTLKLQLPNC 706

Db 659 PQ--FQGLLANLSYTLQDGHMRSGRLPFGDSHELSGNTSITP--DKSCLDFHFPFIC 715

Qy 707 TEDVPSPVLRNFSV---GTPLSAFGN--LRPVLAEADAQLFTALFPFKNCGNDNICQ 762

Db 716 IQDLISPINVSNLFSLEEGTPRDQKGRAMQPIILRPSIHTV--TKEI PFKNCGEDKKCE 774

Qy 763 DDLSTTFPMSLDCLVVGSP-----REFNVTVTVRNDGEDSYRTQVTFPPFLDLSYKY 816


```
Db 775 ANLTSSPARS-----GPLRLMSASLAVETWLTNSCEDAYVVRDLDDPFRGLSFRKV 827
Qy 817 STIQORSQSWRLACESASSTEVSAL--KSTSCSINHSPENSHVNTNITFDVDSKAS 875
Db 828 EMLQ-----PHSRMPVSCBEL--TEGSLLTXTLKNVSSPIFKAGQSVQVMFNLTLLNS 882
Qy 876 LGNKLLKANVTSEN-NMPTNKTEPQLELPVKYAVYVMTVSHGVSTKYNFTASENISR 934
Db 883 WEDFVELNGVHCENENSSLOEDNSAATHIPVLPVNIUTKQENSTLYISPTPKGPKTQ 942
Qy 935 VMOHQOVNGLGORSIPISLVELVPRVLTQVWDRPQ-----VTPSENLS-----TCHTK 986
Db 943 QVQHVTVQ-----RIQPSAYDRNMPT-LEALGVPRPHSEDLITYTWSVOTDPLVTCHE 996
Qy 987 E-RLPSSDFLAELRKAAPVNVCSIAVCQRIQCDIPFGQEBFNAFLKGNLSFDFWIKTS 1045
Db 997 DLKRPSS-----SEAPCLPGV--QPRCPIVE--RWEILQVGTVELSKERAS 1041
Qy 1046 HNHLLIVSTABILFNDSVFTLLPGQAFVRSQTEKVEFEFVNPPLPLIVSSVGLLL 1105
Db 1042 -STLSLCSLSVSFNSSKHFLYGSKA-SEAQVLVKVDLIHEKEMLHVYVLSGIGLVLL 1099
Qy 1106 ALITAALYKLGPPKROYKDMW-SEGPPGCAEP 1136
Db 1100 FLIFLALYKVGPKMLKEMKEMADGGVNGSP 1131

RESULT 9
Q98TF1 PRELIMINARY, PRT; 1196 AA.
AC Q98TF1;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE CD11-1.
GN CIA1.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Nakao M., Miura C., Fujiki K., Yano T.;
RT "Molecular cloning of a leukocyte integrin from the common carp.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB048536; BAB39134.1; -.
DR HSSP; P20701; 1LFA.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001969; Asprotease_AS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; WVF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00032; wva; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; WVFADOMAIN.
DR SMART; SMC0191; Int_alpha; 5.
DR SMART; SMC0327; WVA; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50234; WVFA; 1.
SQ SEQUENCE 1196 AA, 132477 MW, 9369C807E7DCA53B CRC64;
Query Match 24.0%; Score 1409; DB 13; Length 1196;
Best Local Similarity 31.5%; Pred. No. 5e-95;
Matches 375; Conservative 224; Mismatches 447; Indels 144; Gaps 39;
Qy 1 ENLDTENAMTFQENARG-FQGSVVQLQ-GSR--VVVGAQEIIVAAQRGSLQCDYSTGS 56
```

```
Db 32 FNIDTEPLRPNGAPEDFFGVSVYQTEFGNRKQIIVGAPLE---GNSGEMYSCTADLQ 88
Qy 57 CEPTRLOVP-----VEAVNMSGLSLAATTSPPOLLAGPTVHOTCSENTYVYKGLCFLRGS 112
Db 89 CQ--RLQRPGBSESVRFGMSAAVSSAALTS-----CSYPFHEDCDGNSYLNGVCYQPS 140
Qy 113 NLROQPOKFPBALRCPQEDSDIAFLVDGSGSIIPHDFRRAKEFIISTVMBOLKSKSLTFS 172
Db 141 SL-QAVGNFTAAQBCSKREVLNVLFDGSSSMKTVDFEMKNKFIKDIKMKLSNSSIKFA 199
Qy 173 LMOYSEPRTHFTKEPQNNPNPSLIKPIITOLLGRHTATGIRKVVRELN--ITNGARK 231
Db 200 AVQSTIDVTVDNDYQSGSAEKLAKB--THMKSLNTHKADYILKLNLSLMSGADS 258
Qy 232 NAFKILILITDGEKFGDPLGYED--VIPEADREGVIRYVIGVGDAFASHRSKRSQELNIVAS 289
Db 259 KAQKALVIITD---GDPDNDYVNLKCKDEQNILAVIIGV---KVDLIBLTOLAS 309
Qy 290 KPPRDHVPQINNFAALXTIIOQLREKIFAIGTGTGSSSSSEHEMSQEGRSAAITSNGL 349
Db 310 EPEKNFTYIKDYSGLKLNLNLOKTYNIBGSYDAQGRDRQKLSQSGFSVYQEBBSVI 369
Qy 350 LSTVGSYDWAQGVFLYTSKEKSTETINMTRVDSDMN--DAYLGYAAAIILNRVQSILVCLAP 408
Db 370 VGSVGSNDWRGALYEVTC--SGSDPKETEIIDPAVNDKSYMGYSTVLGHRGVSLLFCAP 428
Qy 409 RYQHIGLVAMPFRONTGMWESNANVKGTQIGAYFGASLCSDVDNGSTDLVLIGAPHYE 468
Db 429 RAETHGLVTLFTKNTQNTVTVTSNINEQIGSYFGASLLDLDVSDGSDFTLVGAPLTYQ 488
Qy 469 QTRGQSVSVCPLPRGORARWQCDVLYGEO-----GQPWGRFGAALTVLGVNDGDKLTDV 523
Db 489 SQ-----PRTEGLRVYVLSLSEQKYFOKTLNVYSQSTTGEPAAVASLKDNGDGLSDV 540
Qy 524 AIGAPGEDNREGAVYLPHTGTSGSGISPSHS--QRIAGSKLSPLQYFGOSLGGQDLNDG 582
Db 541 AVGAPLE--NEGTVVILGDRTHGINPELTQRISSVQSIVLPGAQFGVSLTGQMDMNDN 598
Qy 583 LVDLTVAGQGHVLLRQPVLRKALMEFNPREVARNVFECDQVVKGEAGEVRVCLHV 642
Db 599 LTDIVIGAQGGIVLLKARPVMSVAQSUSPKEISLNTFECPGS--NAFNAFLTSCFTV 656
Qy 643 OKSTRDLREGQIQSV--VTYDLDALDSCRPHSRVAFNETKNSTR--ROTQVLGLTQTCET 698
Db 657 TERISS---TGSLEKLNVLNLDVVRGMSRGFFDQSSVSSRTLOQSVLLDSSGSCFN 713
Qy 699 LKLQLPNCIEDPVPSPIVLRNLSLVGTPLSFGNLRPVLAEDAQRLLFTALPFEKNGCND 758
Db 714 FSIFMLRCVADTVSPLKIRNFS--QTEMLS-GNSVAVLDVHSRTBEENVVEFFQRCNCSN 770
Qy 759 NICODDLISITPSFMSLDCLVVGGPREFNVTVTVRNDGBDSYRTQVTPFPFLDLSYRKYST 818
Db 771 NSCVADLKLNFSPTN--NTLVVENQAHTVQVSLANPGDDSYNTSIVLHYPGISLSKFPD 829
Qy 819 LQONORSQSWRLACESASSTEVSGALKSTSCSINHPIFPENSEVFTNITPV---DSKAS 875
Db 830 IKPSRTR-----SSCGDRSDGATNRTTCSIDLVPYRSGTTTQFLGTFTFMKWDNDSS 881
Qy 876 LGNKLLKANVTSENMPRTNKTTEFOLELPVKYAVYVMTVSHGV-STKYNFTASENISR 934
Db 882 NMEIMITANDNDNGM---SDTEVRSSVFPQPAVDLAISLVAEDSVTYMNFSLDRGPK 938
Qy 935 VMOHQOVNGLGORSIPISLVELVPR-----LN 963
Db 939 PLNITYKVENSGRKLDPVSVTLTLCQTPHVILATPHTFPMHEVHSFTSSYHQIMCLLN 998
Qy 964 QTVIWRDPOVTFSENLSSTCHTKERLPSSHDFLAELRKAAPVNVCSIAVCORIQ-----C 1017
Db 999 KELLFPSELSAVQVRSTGSLREVEC-SQF--DLNKSVAHNLITADALQNKEYES 1055
Qy 1018 DIPFPGIOEE--FNATLKGNLSFDWYIKTNSHLLIVSTAILFNDSVFTLLPGQAFVR 1075
```

```

Db 1056 KTSYFVFRKDNVFSISAEINLYNTSYNOTS-----SELKYNPH-----R 1094
Qy 1076 SOTETKVEFEVFNELPLIV-GSSVGGLLLLALITAAALYKLGFFKQYKD 1124
Db 1095 SOTEVKVE-FVVPSSMLIVCTGAVGGFFIILFLLLKCGFFKRNRPD 1143

RESULT 10
Q96HB1
ID Q96HB1 PRELIMINARY; PRT; 1086 AA.
AC Q96HB1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008777; AAH08777.1; -.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Hypothetical protein.
SQ SEQUENCE 1086 AA; 119223 MW; F6PF2546E8C632P9 CRC64;

Query Match 23.1%; Score 1359.5; DB 4; Length 1086;
Best Local Similarity 32.3%; Pred. No. 26-91;
Matches 378; Conservative 185; Mismatches 431; Indels 175; Gaps 37;

Qy 1 ENLDTENATQ--ENARGGQSVVQLGSRVVGARQELVAANORGSLYOCDYSTGSC 58
Db 26 YNLVDRGARSPPPRAGRHFGTRVLQV-GNGVINGAPGE---GNSTGLYQCGSTGHCL 81
Qy 59 PIRLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTYVYKGLCFILGSLRQOP 118
Db 82 PVTLR-GSNYTSKYLGMTLA--TDP----- 103
Qy 119 QKEPALRGCPQEDSDIAFLVDGSGSIIPHDPRAKEPISTVWEQKKSKTLFSLMOYSE 178
Db 104 -----TDGS-----ILFAAVQFST 117
Qy 179 EFRIHFTKEFQNNPNRSLIKPIQLLGRTHATGIRKWKRELPIFNITNGARKNAFKLI 238
Db 118 SYKTEFSDYVYKRDPAKLLKHVKHMLLTTFGAINVATEVEFREELGARPDATKVI 177
Qy 239 LITDGEKGDPLGYEDVTPADREGVIKRVIGVGDAFRSEKSRQELNTVASKPRPDHFQ 298
Db 178 IITDGE--ATDSGNTDAKD-----IIRYIIGIKHFQTKESQETLHKFASKPASEFVKI 230
Qy 299 INNFALKTIONOLREKIFALEGTQTGSSGFHEMSQEGPSAATISNGPLISTVGSYDW 358
Db 231 LDTFEKLDLFTLQKLVIVTEGTSKQDLTSFNMELSSGISADLSRGHAVVGVAGKDW 290
Qy 359 AGGVF-LYTSKEKSTFINNTRVDSMDNDAYLGAAAA-IILNRRVOSLVLGAPRHIGLV 416
Db 291 AGGFGLDKADLQDDTFIGNPELTPTEVRAGYLGYVTWLPSPRQKTSLLASGAPRYQHMGRV 350

```

RESULT 11

```

Q98TF0
ID Q98TF0 PRELIMINARY; PRT; 1187 AA.
AC Q98TF0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CD11-2.
DE CIA2.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Fujiki K., Nakao M.;
RT "Molecular cloning of a leukocyte integrin from the common carp.";

```

```

Qy 417 AMFR--QMTGWESNANVKGTQIGAYFGASLCSDVDNSGSTDVLVIGAPHYFETRCGQ 474
Db 351 LFFQBPQGGHWSQVTHGTQICSYFGEELCGVDVDGDETELLIGAPLYGQRQGR 410
Qy 475 VSVCLPRGQPARWQCDAY--LYGEQCPWRFRGAALTVLGDVNGDKLTVAIGAPGEED 532
Db 411 VFYI-----QRRQLGFEVSELOQDPGYPLGRFGEAITALTIDNGDLVDVAVGAPLEE- 464
Qy 533 NRGAVYLFHGTSGSIGSPHSORLAGSKLSPRYFGQSISGGQDLTMDGLVDLTVGAQG 592
Db 465 -QGAVYIFNGRHG-GLSPQSPQRLEGTVLSGIOMFGRSHGVKDLBEGDLADVAVGAES 522
Qy 593 HVLLRSQPLVRVKAIMEFNPVARNVFNQDQV--KGEAGEVRVCLVHKVSKTRDLRL 651
Db 523 QMIVLSRPVDMVTLMSFSPAEIPVHEVECSYSTSNKMEGVNITICFQI-KSLIPQF- 580
Qy 652 EGOIQSVVYDLDLDCSRPHSRVFNKSTKSTQVLTQVLTQTCETLKQLPNCIEBPV 711
Db 581 QGRIVANLTYLQDCHTRRRGLFPGRHLELRNIAT-TSMSCDTFSPPFPVCVDLI 639
Qy 712 SPIVRLNPSL---VGTPLSAPGN-----LRPVLABDAORLFTALPFPFNKCGNDNICQ 762
Db 640 SPINVSINFLWREGEPTDORAGKDIPPLRPSLHSETWEI-----PFEKNCGEDKKCE 694
Qy 763 DLSITSPMSLDCLVGGPREFNVTYVRNDGEDSYRTQVTFEFLDLVYKVKYSLTQNQ 822
Db 595 ANLRVSFSPAKSRALRUTAFASLSVELSLNLEDAYTWQDLHPPLGLSPRKVEL--- 751
Qy 823 RSQRSWLACES--ASSTEVSGALKKTSCTSINHPIFPENSEVTFNITFDVDSKASLGKL 880
Db 752 KPHSQIPVSCDELPEBSRLLSRAL---SCNVSSPIFRAGHSVALQMMFNTLVNSWGDV 808
Qy 881 LKANTVSENN---MPTNKTPEQLELPVKYAVVMVTVSHGVSTKYNLFTASENTSRVM 936
Db 809 ELHANVTCCNEDSDLENSGATTI---IPIDYINILIQOEDSTLVSVFTPKGPKIHQV 865
Qy 937 QHOYQV---SNLQORSIP-LSLVLVPLVRLMOTVINDRPQVTFSENLSSTCHTK--BRLP 990
Db 866 KMYQVRIQPSIHDHNIPTLEAVGVVQPPSEGIPTHQMSVQMBPPV--PCHYEDLERLP 923
Qy 991 SHSD--FLAELRKAPVWNCSLAVCQRIQCDDIPFGIOEBFNATLKMISFWDYIKTSNH 1048
Db 924 DAEPFCLPGLALFRCPVW-----PRQELVQVIGTLELVGEIEAS--SM 964
Qy 1049 LLIVSTAEILFNDSPVFTLLPGQAFVRSQETKVEPEVEPNPLPLIVGSSVGGLLLALI 1108
Db 965 FSLCSSLSISFNSSKPHLYGSNASL-AQVMKVDVYVYKQMLYLYVLSGIGLLILLI 1023
Qy 1109 TAALYKLGFTKQYKDMMSSEG-GPPGABP 1136
Db 1024 FIVLYKVGFTFRNLKERMEAGRGVPNGIP 1052

```

RL	Submitted (SEP-2000) to the EMBL/GenBank/DBDJ databases.
DR	EMBL; AS048537; BA39135.1; ..
DR	HSSP; P20701; ILFA.
DR	GO; GO:0008305; C:integrin complex; IEA.
DR	GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR	GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR	GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR	InterPro; IPR001969; Asparticase AS.
DR	InterPro; IPR000413; Integrin_alpha.
DR	InterPro; IPR002035; VWF_A.
DR	Pfam; PF01839; FG-GAP; 3.
DR	Pfam; PF00092; vwa; 1.
DR	PRINTS; PR01185; INTEGRINA.
DR	PRINTS; PR00453; VWFADOMAIN.
DR	SMART; SM00191; Int_alpha; 5.
DR	SMART; SM00327; vwa; 1.
DR	PROSITE; PS00141; ASP_PROTEASE; 1.
DR	PROSITE; PS0234; VWFA; 1.
SQ	SEQUENCE 1187 AA; 131778 MW; 85EDCTCA86B1C59 CRC64;
Query Match 23.1%; Score 1358.5; DB 13; Length 1187;	
Best Local Similarity 31.6%; Pred. No. 2.8e-91;	
Matches 377; Conservative 206; Mismatches 452; Indels 157; Gaps 43;	
QY	1 FNLDTENAMTFQENARG-FQGSVVQLQ-GSR--VWVGAQOEIVAAQNGSLYQCDYSTGS 56
DB	32 FNIDTEHLEFNGTPEDFGVSVVTEFGNKKLIIVGAPLE---GNSAGEMTCTADLOS 88
QY	57 CEIRLOVP-----VEAVNSLGLSLAATTFPPQLLACGPTVHCTCSENTVKGCLFLGS 112
DB	89 CK--RLQRPQGSSEVRFPGNSAAVSSAALTS-----CSPYFAHECDGNSYLNGVCYQFNS 140
QY	113 NLRQPKPPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKBFISTVMEQLKSKTLFS 172
DB	141 SL-QAVSNFTAAVQECSEKREVNLFVDFGSSMKAVFDMKNKFIKDVMMKLSNSSIKPA 199
QY	173 LMQVSEFRHPTPEKQNNPNRSLIKPTOLLGRTHATGTRKVVRELFN-ITNGARK 231
DB	200 AVQSTBIRTVDFDNDYQNGSAEELMKR-RHMKSLTNTYKAINVLYKNVLSVSGADP 258
QY	232 NAFKILLITDGERFGDPLGYED--VIPEADREGVIRVIGVDGAFRSEKSRQELNLTAS 289
DB	259 NAQVALVIITD-----GPSUNDDYVILNIQDEQILYIIVGV-----KYDLITLTQLAA 309
QY	290 KPPRDHVQINNFPALKTQNLREKIFAEIGTGTGSSSFEHEMSQEGFSAITNSGPL 349
DB	310 EPKLNNTFYIOEYNGLGLDNLQKTYNIEGSKAEHGRDRQKELSQSGFSVYQEESVI 369
QY	350 LSTVGSVDWAGGVLYT---SKERSTFNTMTVDSDMNDVGLGYAAAILRNVRVQSLVLG 406
DB	370 VGSVGSNDWRGALYEVWVGSGSKFQTEITDPAVN---KDSFMGISTVLGKRGHGVSLPSG 426
QY	407 APYQHIGLVAMFRQNTGMESNANVKGQTQIGAYFGASLCSDVDVDSNGSTDVLVIGAPHY 466
DB	427 APRAHTGLVTLPTKNBSTWTVMENINGEIGSYFGASLSLDDVSDGSDDFLLVAGPLF 486
QY	467 YE-QTRG-GQVSVCLPFGQPARWCDAVLYGEQ--GQPWGRFGAALTVLGDVNGDKLTD 522
DB	487 YCSQFRAEGRLYVYTL-----SEQYSQRTLOSTTGRFATSLASLKDNLGDLSD 535
QY	523 VAIGAPGEDNRGAVLPHGTSGGSGSPSHS-QRIAGSKLSPRLQYFCQSLSGQDITMD 581
DB	536 VAVGAPLE--NEGVMYIIVGLDGTGHNPEHAPQRIPARSVLPGLQFGVSLSGQMDMND 593
QY	582 GLVLDLVGAQGHVLLRSQPLRVKVAIMEFNPREVARNVFECDQVVKGEAGEVRVCLH 641
DB	594 NLPDVIQTGGVILLNARPNVMSYSAQLSFNPMISLNYFPCPGS--NAFNAFNLITSCFT 651
QY	642 VOKSPFRLRQIQGV--VTVDLALDSGRPHSRAVFNETKNSTR--RQTVGLGLTQCE 697
DB	652 VTERISS---TGSLEKLVNSLNLNDVVRGKSRGFFDPMDSSRTLQOSVILLDSGSSCS 708
QY	698 TLKQLPNCIEDPVSPIVLRNFS---LVGTPLSAFCGNLRPVLAEQAQRLLFTALFPPEK 753
DB	709 NFSIFMLECVADTVSPLKIRNFSQTKVLSGNLAVL-DIQSRTSEYVEVL-----PQR 761
QY	754 NCGNDNICQDDLSITFSFMSLDCLVWGGPREFNTVTVVRNDGDSYRTQVTFPPDLDSY 813
DB	762 NC-NSNSCVADKLNFSTN-DTLVVENQAHTVTVLSLANFGDDSYNTSIVLHYPEGUSL 819
QY	814 RKVSTLQNRORSRWLACBSASSTEVSGALKSTSCSINHPIFFPENSEVTNITPDV--- 870
DB	820 SKFDAIKPSRTR-----SSCGDRDSGATNRTCSINLFPVRSGLTTQFLGTFPRVTKW 871
QY	871 DSKASLGNKLLKANVTSENMPRTNKTFFOLELPVKYAVVWVTSYSHGV-STKYLNFAS 929
DB	872 DYDWSDRMERTITANSNNNGNM---SDMSVRSIPQVFAVELAISLVAEDSVTYLNFSL 928
QY	930 ENTSRVWQHGYOVNLSGORSIPISLAVPLVPVRLAQTVIWDROPQVTESENLSTCHTKERL 989
DB	929 DRGPKPLNIIYKVNHLGKLPVSVTLSPCQ-----TTHVTLTPENFSMQ 974
QY	990 PSHSDFLAELRKAPVVC-----SIAVCQRIQCDIPFFGIQEE 1027
DB	975 EVHHSFISSYHQ--IIMCLLNKHLFPSPSLSAVQTRKTRGSLWVC-----VSSISTGBI 1026
QY	1028 FNATLKGNI-----SFDWYIKTSHNLLIVSTAELFNDSVPTLLPQ-OCA 1072
DB	1027 FRSSV--NLMAEAVLQNVKEYESKYSFY-EFRDRHVFNIS-AELAFNTSRYNQSTGLKYN 1082
QY	1073 FVRSQTEVKVEPPEVNPPLIVGSSVGGLLLLALITAAVLYKLGFFKQYKD 1124
DB	1083 PHRSQTEVKVEPVPFPPSRMLIVCTGAVGSGFPPLIILILLKCGFFKGRPD 1134
RESULT 12	
Q8HZVO PRELIMINARY; PRT; 927 AA.	
AC	Q8HZVO;
DT	01-WAR-2003 (TREMELrel. 23, Created)
DT	01-WAR-2003 (TREMELrel. 23, Last sequence update)
DT	01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE	Lymphocyte function-associated antigen 1 (Fragment).
OS	Bos taurus (Bovine).
OC	Rukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC	Bovidae; Bovinae; Bos.
OX	KCEI_TaxID=9913;
RN	[1]
RA	SEQUENCE FROM N.A.
RA	Thumbikat P., Kannan M.S., Maheswaran S.K.;
RT	"Sequence of the alpha subunit of bovine lymphocyte function-
RT	associated antigen 1."
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBDJ databases.
DR	EMBL; A2440778; AAN63636.1; ..
DR	PIR; A32039; A32039.
DR	GO; GO:0008305; C:integrin complex; IEA.
DR	GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR	GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR	InterPro; IPR000413; Integrin_alpha.
DR	InterPro; IPR002035; VWF_A.
DR	Pfam; PF01839; FG-GAP; 3.
DR	Pfam; PF00092; vwa; 1.
DR	PRINTS; PR01185; INTEGRINA.
DR	SMART; SM00191; Int_alpha; 4.
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR	PROSITE; PS0234; VWFA; 1.
FT	NON_TER 1
FT	NON_TER 927
SQ	SEQUENCE 927 AA; 102523 MW; 02E2CF09917081EC CRC64;
Query Match 21.7%; Score 1278; DB 6; Length 927;	
Best Local Similarity 34.9%; Pred. No. 1.7e-85;	
Matches 341; Conservative 166; Mismatches 385; Indels 86; Gaps 28;	

AC O88340; 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Integrin alpha E1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Rodentia; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
SEQUENCE FROM N.A.
MEDLINE=98056820; PubMed=9394838;
Brenan M., Rees D.J.G.,
RT Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020045; AAC23662.1; .
DR HSSP; P11215; 1BHQ
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VMP_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VMPADOMAIN.
DR SMART; SMO0191; Int alpha; 4.
DR SMART; SMO0327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VMPA; 1.
SQ SEQUENCE 1167 AA; 128970 MW; D88A2C38ACDC2AAB CRC64;
Query Match 19.9%; Score 1167.5; DB 11; Length 1167;
Best Local Similarity 29.0%; Pred. No. 4e-77;
Matches 358; Conservative 214; Mismatches 456; Indels 205; Gaps 43;
QY 1 FNLDTENA--MTFOENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNGS-----LYQCDY 52
DB 20 FMDVDVAVVATLQPGAPAVLSLLHLDPSN-----NQTCLLVARRSSNNTAALYRCAL 74
QY 53 STGSCBPIRLQVPVEAVNMSLGLSLAAT--SPQLLAC-GPTVHTQCSNTYVKGICFL 109
DB 75 SI-SPDSIACQ-PVEHICMPKRGYQVTLVGNHNGVLVCIQVQARKFRSLNSELTGACSL 132
QY 110 FGSNLRQOPQKFPPEALRG-----C-----POE 131
DB 133 LTPNLDLQAOQAFSDLSGLDPLDGAHVDSGDYCRSKGSGTGEKKASARRRTVEDEED 192
QY 132 DSDIAFLVDGSGSIIPHDFFRAKEFIIVMEQL--KSKTLFSLAQYSEEFRIHFTKCF 189
DB 193 GTEIAVLVDGSGSIEPSDFQKAKNFISTMKNFYKCFECNFALVQYGAIVQTEFDLQES 252
QY 190 QNNPNRSLKPIIQLLGRHTATGIRKVVRELNFNIGARKNAFKILILITDGEKCDP 249
DB 253 KDINASTAKVQSIYQVKEVTKTASAMQHLNDNIFIPSRGSKKALKVWVVLTDGDI 312
QY 250 LGYEDVTPEDAREGVIRVIGVDAPFRSEKSRQELNTVASKPPRDHVFQINNFEALTIQ 309
DB 313 LNLITVINSFQWQGVTRPAICGVDAFKNNNTYRELKLIASDPKEAHTFKVTVNSALDGL 372
QY 310 NOLREKIPAIETGTGSSSSPEHEMSQEGFSAATTSNGP-LLSTVGSYDWAAGGVLY-TS 367
DB 373 SKLQORIWHMEGT---VGDALQYQLAQTFSAQIDKRGQVLLGTGVAFNWSGGALLYSTQ 429
QY 368 KEKSTFENMT-RVDS-DMNDAYLVGAALILRNVRQSLVLCAPRYOHIGLVAMFRONTGM 425
DB 430 NERGCFNQAKESRTVOYSYLVSLAVLKAHGISYVAGAPRHKLGAIVELKEDR- 488
QY 426 WESNA---NVAGTQIGAYFGASLCSVDVDSNGSDTLVIGAPHYYEYQTRGGQSVCPILPR 482
PRT; 1167 AA.

174 MOYSEFRIHTFKBFONPNRSLIKPIITOLLGRHTATGIRKVVRELNFNTGARKNA 233
DB 4 VQPSYFRTFTFLDYIQKDPDALLAGVKGMRLLTNTFGALNVVKEVFPDLGAPDA 63
QY 234 PKILLITDGEKFGDPLGYEDVTPADREGVIRVYVGVDAFRSEKSRQELNTVASKPPR 293
DB 64 TKVLIIITDGG---PPTNTTLMRPTSS-----RSLIGIGKFKFKESQALHQFASKPVE 116
QY 294 DHVFOINFEALTKITQNLREKIPAIETGTGSSSSPEHEMSQEGFSAATTSNGPLISTV 353
DB 117 EPVKLIDTFEKLKDLITELQKLIYVIGTSGKDLTSFNMELSSGISADLSEGHGVGAV 176
QY 354 GSYDWAGVVF-LYTSKSKSTFNMTVRVDSMDNDAYLVGAAL-IILNRVQSLVIGAPRYQ 411
DB 177 GAKDWAGGFLDKADLKSTFTVNEPLTAVESRAGVLYTVTLPSRGTGMSLLATGAPRYQ 236
QY 412 HIGLVAMPRO--NTGWSNANVKGTQJGAYFGASLCSVDVDSNGSDTLVIGAPHYYEQ 469
DB 237 HVGRVLLFQPKRGFGFWQIQIDIGIQISTFGBELCGVDVDRDBETELLIAAPLYICE 296
QY 470 TRGGQVSVCPILPRGQARWQCDVAVLYGEGQPGWGFGAALTVLGVGVNDGKLTDAVGAIGP 529
DB 297 QRGGRVFTV---QKIQLEFQWVSELOGETGYPLGRFGAIAAALTIDINGDELTDVAVGAPL 353
QY 530 BEDNRGANVLFHGTSGSGISPSHSORIASGKSLSPLOYGQSLGCGDLTDGGLVLTVG 589
DB 354 EE--QGVYIFNGQQG-GLSPRPSORIEGTQFSGIQMFGSRHGVKDLGGGLADVAVG 410
QY 590 AQCHVLLARSQPLRVKATMBENPFEVARNPECDNVVKGKEAG-EVRVCLRHQK--ST 646
DB 411 ABGQVILSSRPVDIITVSFSPAEPVHEVECSYSTSNQKKEGVNLTVCFQVKSLLIST 470
QY 647 RDRLEGQIQSVVTDALDSCGRPSRAVFNETKSTRQTVGLTGTCTCLKQLDENC 706
DB 471 ----FQGHVAVNLVTLQDGHRTSRGLFPQGRKHLIGNTAVTPV-KSCFVFMFHPFIC 525
QY 707 IEDPVSPTVLRNLSL---VGTPLS--AFGNLRPVLAEADAQRLFTALPPEKNGCQNDNIC 761
DB 526 IODLSPINVSLSYSLWEEGTTPRDLRDRDIPILKPSPHLETKETPFEPKNCQEDKNC 585
QY 762 QDDLSTFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTFPPDLDSLVRKYSTQON 821
DB 586 EADLKLAFSDMRSKIIRLTPSASLSVRLTRNTAEDAYVQVTLSPQGLSPKRVKIL-- 643
QY 822 QRSQSRWLACASASTEVSGALKSTCSINHPFENSEVTENITFDVDSKASIGNKLL 881
DB 644 -KPHSHVPVGCCELPEEAHVHS-RALSCNVSPFIFGDSMDIQMNFNLQKSGWGFIE 701
QY 882 LKANVTS-----ENNMPTNKTETFOLELPVKYAVVWVTSHGVSSTKYNLFTASENTS 933
DB 702 LQANVSCNEDSSLLSDNSATTS-----IPWYPIVNLTKDQENSTLYISFTPKSPRI 754
QY 934 RVWQHOYQV---SNLQORSLSISLVFLVPLRLNQTVI---WD---RQVTFES-ENLST 982
DB 755 HHVKHIYQVRIQPSNYDNMP-FLEALVRVPRVHSEGLITHKWSIQMEPPVNCSPRLNESP 813
QY 983 CHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDIPFGIQEBFNATLKNLSFDWYI 1042
DB 814 SDEAE-----SCSFGI--EFRCFIDF---RQELLVQVNGMVELRGTI 850
QY 1043 KTSNHLILVSTABILFNDVSTPLLPQOGAFVRSQETETKVEFPFVNPPLPIVSGSVGL 1102
DB 851 KAS-SMLSCSLAISFNSSKHFLHGRNASM-AQWVKVLDLVYKEMLYLVLSGIQGL 908
QY 1103 LLLALITALYKLGFEKR 1120
DB 909 LLLFLIFIALYKVGFEKR 926
PRELIMINARY; PRT; 1167 AA.

RESULT 13
O88340
ID O88340

Db 489 -BEDAFVRRIGRQMSYFGSVLCPVDIDMDGTTDFLLVAAPYHIRGREGVYVQVPE 547

Qy 483 GQARWQCDVLYGEGQOPWGRFGAALTVLGDVNGDKLTDVAIGAP-----GEEDNRGA 536

Db 548 -QDASFSLAHTLSHPLGNTSRFGFMAAVGDINQDKFTDVAIGAPLEGFAGDGASYS 606

Qy 537 VYLFPHGTSGGSIIPSHSQRTAGSKLSPLOYFQOGLSGGDLTMDGLVLTWCAQGHVLL 596

Db 607 VYIYNGHSG-GLYDPSQIRASSVAGSLHYFGSVSGGLDFNGDGLADITVGRSDAVV 665

Qy 597 LRQOPVLRKAIMFEPREVARNVFECNDQVVGKEAGEVRVCLHVQKS-----TRDLREG 653

Db 666 LRARPVVDLTVSMTFP-----DALPWFVIGKM--DVNLCFEVDSSVASEPGLREM 715

Qy 654 QIOSVVTYDIALDGRPHSRVAVNETKNSRTROTQVGLTQTC-----696

Db 716 FLNFTVDVY-----TKQRQLACEDSSGQCLRWKNGSFCEHFWLI 760

Qy 697 ETLLQLPNCIEDPVSPIVLRNLFSLVGTPLSAFQNLN-----PVLAEADAQRLTALF--P 750

Db 761 STEEL-----CEBDCFSNITIKYVE-----PQISGGERDYENPL--DRYKESALFQLP 809

Qy 751 FEKNCGNDNTQQDLSITTFMSLDCLVGGPREFNTVTVRNDGEDSYRTQTTFPFLD 810

Db 810 YEKDCKNVFCIABIQLTIN--ISQELVGVTVKEVTMNIISLTNSGSDSYMTNMAIYPRN 868

Qy 811 LSYRKVSTLQNSORSWRRLACESASSTEVSGALKSTSCSINHPIPFENSEVTENITFDV 870

Db 869 LQFKKI-----QKVPSPVQCDPKV-----ASVLWNCIKLGHPIIL--KRSVNVSVTWQL 918

Qy 871 DSKASLGNKLLKANVTSENMPNTKTEFQLELPKYAYVMVTVSHGVSTKYLNFTASE 930

Db 919 EESVFPNRTADITVIGSNEKSLARETR--SLOPRHAFIAVLSR--PSVMYMN--TSQ 971

Qy 931 NTSVMQHOVQVNLGORSPLISLAVFLVPLVRLNQTVIWDPRQVTFSENLST-----CHT 985

Db 972 SPSPDKHFFNVHGENLFGAVFQIQICVPIKLQF-----QIVRVKNTKTQDHECTQ 1025

Qy 986 KERLPESHSDPLAEIRKAPVNVCSIAVCQRIQCDIPFEGIOEBFNATLKNLSDWIKTS 1045

Db 1026 SQEPACGSDPVQVHKWHSVVCAL-----TSKNENVTVAAEISVG 1065

Qy 1046 HNHLIVSTA-----BILFNDSVFTLLPGGAFVRQSTKTEPF-----EVENPLPLIV 1095

Db 1066 HTKQLLRDVSLEPLTIGSEIFNKSLEYGLNAB-----NHRKTIVIFLKEETRSPLLI 1119

Qy 1096 GSSVGGILLALITAAALYKLGFFKQVQKMMSE 1128

Db 1120 GSSIGGLLVVILAILFKCGFFKRYQQLNLE 1152

RESULT 14

Q88341 ID O88341 PRELIMINARY; PRT; 1167 AA.

AC O88341;

DT 01-NOV-1998 (TRENDELrel. 08, Created)

DT 01-NOV-1998 (TRENDELrel. 08, Last sequence update)

DT 01-JUN-2003 (TRENDELrel. 24, Last annotation update)

DE Integrin alpha E2 (Fragment).

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OK NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98056820; PubMed=9394838;

RA Brenan M., Rees D.J.G.;

RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:

RT tissue expression reveals phenotypic similarities between

RT intraepithelial lymphocytes and dendritic cells in lymph.;"

RL Eur. J. Immunol. 27:3070-3079(1997).

DR EMBL; AF020046; AAC23663.1; -.

DR HSSP; P11215; 1BHQ.

DR GO; GO:0008305; C:integrin complex; IEA.

DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.

DR GO; GO:0007160; P:cell-matrix adhesion; IEA.

DR InterPro; IPR000413; Integrin_alpha.

DR InterPro; IPR002035; VWF A.

DR Pfam; PF01839; FG-GAP; 3.

DR Pfam; PF00357; Integrin_A; 1.

DR Pfam; PF00092; vwa; 1.

DR PRINTS; PRO1185; INTEGRINA.

DR PRINTS; PRO0453; VWFADOMAIN.

DR SMART; SM00191; Int_alpha; 4.

DR SMART; SM00327; VWA; 1.

DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.

DR PROSITE; PS00234; VWFA; 1.

FT NON_TER

SQ SEQUENCE 1167 AA; 128593 MW; E3FED7E18B5CBBBC CRC64;

Query Match 19.1%; Score 1123; DB 11; Length 1167;

Best Local Similarity 27.7%; Pred. No. 8e-74; Indels 244; Gaps 37;

Matches 334; Conservative 207; Mismatches 421;

Qy 47 LYCCDYSTGSCBPIRLQVPEAVNMSLGLSLAATTPSPOLLACGPTVH-----94

Db 67 LIQC-----ATSIIDEICQCPVEHILMPKRGYQVT 98

Qy 95 -----QTCSENVYVKGLCFLFGSNLQOQPOKPFPEALRG-----127

Db 99 LVRNHNGVLVCTIQVQSRKPRSLNSELTGACSLTLPNLDLQAAQYFSDLEGLVDLGASVNS 158

Qy 128 -----CPQED-----SDIAFLVDGSGSIIPHPFRRAKEPI 157

Db 159 GYVYSGKSGSTGEETYSARRLRQARVEEDEBEAGTEAIVLDGSGSIEPSDFQAKDFI 218

Qy 158 STVMEQL--KSKTLFLSLMOYSEERIHTPFKEFQNNPNRSLIPIPTQLLQRTHTATGI 215

Db 219 STWNRNFEYKEKPECNFAVQYGVQIQTDFDLDSRDINASLAKVQSIQVQKEVTKTASAM 278

Qy 216 RKVVRLEFNITGARKNAKILILITDGEKGDPLGYEDVPEADREGVIRVIVGVGDAF 275

Db 279 QHVLNDIFIPSGSRGALKWVWITDGDIFRDPNLNLTIVISSRMOQGVVRFALGVGNAF 338

Qy 276 RSEKSRQELNTVASKPRDRHVFQINNFEAKTQNOLREKIFAIEGTQTGSSSSPHEMS 335

Db 339 ENNTYRELKLIASDPKAAHTFKVTNYSALDGLLSKMQRIIHMEGT---VGTLYQYLA 395

Qy 336 QEGFSAITSNGP-ILSTVSGYDMAGGVLY--TSKEKSTFINNTRVDSMDNDA---YLG 390

Db 396 QTGFSAQILDKGOVLLGTGVAFNWSGGALLTYNONGRFLNQT--AKEDFRAAQYSLGY 454

Qy 391 AAIILNRNVQSLVLCAPRYQYHIGLVAMFRONTGWESNANVKGTQIGAYFGASLCSVDV 450

Db 455 SVALHKAHGVSVACAPRKLKRGAVFELQKEDGETFWRIEGEQMGSYFGSVLCPVDI 514

Qy 451 DNGSTDLVLIGAPHYTYETROGVSVCPFLPRQARWQCDVLYGECQCPWGRFGAALT 510

Db 515 NMDGITDFLLVAAPFYHIRGEGRVYVYRV--HQDAPFSLVYTLGYGLTSSRFQFAMA 573

Qy 511 VLGDVNGDKLTDVAIGAP-----GEEDNRGAVYLPFGTSGSGISPSHSQRIAGSKLSPR 564

Db 574 AVGDINQDFTDVAIGAPLEGFAGDGASGVYIYNGHSG-GLHASPQCIQIRASSVALG 632

Qy 565 LQYFGQSLGSGQDLTMDGLVLDLTVGAGHVLRLSOPVLRVKALMEFNFEVARNVPECN 624

Db 633 LYYFGMSVSGGLDFSGDDLADITVGSQDVAVVLRSEPPVDLTVSMFTFDALP-----695

Qy 625 DVVYKGEAGEVEVCLHVQKS---TRDLREGQIQSWVTYDIALDGRPHSRVAVNETKN 681

Db 686 ---MAFKQMDVELCPKVDSSAVSPSPGLRGHSLNFTVDVY-----TKQ 727

Qy 682 STRRQ-----TQVGLGTQTCETLKLQLPNCIEDPVSPIVLRNFS 721

Db 728 KQRLQCADRSCCQCLMKWNGSGSLCEHFLISTEBEL-----CBDDCFSNITIKVSYE 781

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:54:39 ; Search time 49.4608 Seconds
(without alignments)
6495.175 Million cell updates/sec

Title: US-09-902-481B-5

Perfect score: 5876

Sequence: 1 FNLDTENAMTQENARGFGQ.....PKRYQXMMSEGGPGAEQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5862	99.8	1153	2 AAW65090	AAW65090 Human Bet
2	5862	99.8	1153	2 AAB07360	AAB07360 Human CD1
3	5862	99.8	1153	5 AAU80252	AAU80252 Human Int
4	5862	99.8	1153	5 ABG61469	ABG61469 Human Bet
5	5862	99.8	1153	5 AA014428	AA014428 Integrin
6	5862	99.8	1153	7 ADP25615	ADP25615 Binding d
7	5852	99.6	1153	2 AAR04136	AAR04136 Alpha sub
8	3460	58.9	1163	2 AAR07120	AAR07120 p150.95 a
9	3446	58.6	1163	2 AAW65091	AAW65091 Human Bet
10	3446	58.6	1163	3 AAB07361	AAB07361 Human CD1
11	3446	58.6	1163	5 ABG61470	ABG61470 Human Bet
12	3444	58.0	1163	6 ABU07406	ABU07406 Protein d
13	3411	58.0	1161	2 AAR78166	AAR78166 Human bet
14	3411	58.0	1161	2 AAW23049	AAW23049 Human bet
15	3411	58.0	1161	2 AAW57491	AAW57491 Human bet
16	3411	58.0	1161	2 AAW65089	AAW65089 Human Bet
17	3411	58.0	1161	2 AAW72825	AAW72825 Human alp
18	3411	58.0	1161	2 AAW73342	AAW73342 Human alp
19	3411	58.0	1161	3 AAB07359	AAB07359 Human alp
20	3411	58.0	1161	5 ABG61468	ABG61468 Human Bet
21	3395.5	57.8	1161	2 AAW23064	AAW23064 Human bet
22	3395.5	57.8	1161	2 AAW65106	AAW65106 Human Bet
23	3395.5	57.8	1161	2 AAW72837	AAW72837 Human alp
24	3395.5	57.8	1161	2 AAW73343	AAW73343 Human alp
25	3395.5	57.8	1161	3 AAB07376	AAB07376 Human alp

26	3395.5	57.8	1161	5 ABG61485	ABG61485 Human Bet
27	3234.5	55.0	1161	2 AAR78169	AAR78169 Rat alpha
28	3234.5	55.0	1161	2 AAW23062	AAW23062 Rat beta
29	3232.5	55.0	1161	2 AAW60004	AAW60004 Rat alpha
30	3232.5	55.0	1161	2 AAW72824	AAW72824 Rat alpha
31	3232.5	55.0	1161	3 AAB07374	AAB07374 Rat alpha
32	3232.5	55.0	1161	5 ABG61483	ABG61483 Rat beta
33	3225.5	54.9	1161	2 AAW65104	AAW65104 Rat beta
34	3225.5	54.9	1161	2 AAW73345	AAW73345 Rat alpha
35	3222	54.8	1161	2 AAW23061	AAW23061 Mouse bet
36	3222	54.8	1161	2 AAW60003	AAW60003 Mouse alp
37	3222	54.8	1161	2 AAW65103	AAW65103 Mouse bet
38	3222	54.8	1161	2 AAW72836	AAW72836 Mouse alp
39	3222	54.8	1161	2 AAB073347	AAB073347 Mouse alp
40	3222	54.8	1161	3 AAB07373	AAW07373 Mouse alp
41	3222	54.8	1161	5 ABG61482	ABG61482 Mouse Bet
42	3218	54.8	1161	2 AAR78168	AAR78168 Mouse alp
43	3211.5	54.7	1151	2 AAW23059	AAW23059 Rat beta
44	3211.5	54.7	1151	2 AAW60001	AAW60001 Rat alpha
45	3211.5	54.7	1151	2 AAW65101	AAW65101 Rat beta-

ALIGNMENTS

RESULT 1

AAW65090
ID AAW65090 standard; protein; 1153 AA.

XX AC AAW65090;

XX DT 28-SEP-1998 (first entry)

XX DB Human Beta-integrin CD11b subunit protein.

XX KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
XX KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
XX KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;
XX KW rheumatoid arthritis.

XX OS Homo sapiens.

XX FN US5728533-A.

XX PD 17-MAR-1998.

XX PF 07-JUN-1995; 95US-00485618.

XX PR 23-DEC-1993; 93US-00173497.

XX PR 05-AUG-1994; 94US-00286889.

XX PR 21-DEC-1994; 94US-00362652.

XX PA (ICOS-) ICOS CORP.

XX PI Van Der Vieren M, Gallatin WM;

XX DR WPI; 1998-206565/18.

XX PT Screening assay for modulators of integrin binding - using immobilised or
XX PT labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.

XX PS Example 5; Fig 1A-D; 106pp; English.

XX CC This sequence represents a human beta-integrin CD11b subunit which is
XX CC used to describe a method for identifying compounds that modulate the
XX CC interaction of the beta-integrin alpha-d subunit with a binding partner
XX CC of alpha-d which involves contacting an alpha-d polypeptide with an alpha
XX CC -d binding partner, one of which is immobilised and the other of which is
XX CC labelled, in the presence of a test compound, and determining if the
XX CC compound affects binding between the alpha-d polypeptide and alpha-d
XX CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment
XX CC comprising the cytoplasmic, transmembrane or extracellular domain of
XX CC alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC and rheumatoid arthritis
XX
SQ Sequence 1153 AA;

```
Query Match          99.8%; Score 5862; DB 2; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0
```

Qy 61 RLQVPEAVNMSLGLSLAATTSPPLLACGPTVHQTCSENTYVKGICFLFGSNLRQOPQK 120
 |||||
 Db 77 RLQVPEAVNMSLGLSLAATTSPPLLACGPTVHQTCSENTYVKGICFLFGSNLRQOPQK 136

Qy	121 FPEALGCGPQEDSDIAFLVDGSGSI PHDFRRAKEPISTVNEQLKSKTFLSLMOYSEEP 180 : : : : : : : : : : : :
Db	137 FPEALGCGPQEDSDIAFLVDGSGSI PHDFRMKEFVSTVNEQLKSKTFLSLMOYSEEP 196 : : : : : : :

Qy	181	RHFTFKBFQNNPNRSLKPIQTLLGRTHTATGIRKWVRLFNITNGARKNAFKILILI	240
Db	197	RHFTFKBFQNNPNRSLVKPITQLLGRTHTATGIRKWVRLFNITNGARKNAFKILIVI	256

Qy	Db
241	257

Qy	301	NFEALKTIONQLREKIFAI	EGTQTSSSSS	PEHEMSQEGFSA	IT	NGPL	ST	VG	SYD	WAG	360
Db	317	NFEALKTIONQLREKIFAI <th>EGTQTSSSSS</th> <th>PEHEMSQEGFSA</th> <th>IT</th> <th>NGPL</th> <th>ST</th> <th>VG</th> <th>SYD</th> <th>WAG</th> <th>376</th>	EGTQTSSSSS	PEHEMSQEGFSA	IT	NGPL	ST	VG	SYD	WAG	376

[illegible]

Qy 421 QNTGWESNANVKGTIQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTGGQSVCP 480
 Dd 437 QNTGWESNANVKGTIQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTGGQSVCP 496

Qy 481 PRGQARWQCDAVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 540
|||
Db 497 PRGQARWQCDAVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 556

Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLOQYFGQSLSGGQDLTMDGLVDLTWGAGQHVLLLRSQ 600

Db 557 HGTSGSGISPSHSQRIAGSKLSPRLOQYFGQSLSGGQDLTMDGLVDLTWGAGQHVLLLRSQ 616

QY 601 PVLRVKAIINEFNPREVARNVFECNDQVWKGREAGEYRVCLHVKQKSTRDLREGQIQSVVT 666
 |||||
 Db 617 PVLRVKAIINEFNPREVARNVFECNDQVWKGREAGEYRVCLHVKQKSTRDLREGQIQSVWT 676

QY 661 IDALDGRGPHSKAVFNEIANSIRKQIVDGLIQLCBILAQDFENCIEDFVSFIVLRNF 7366

D6 677 YDLALDGRGPHSKAVFNETKNSTRRTQVGLGLTQTCTLKLQPNCIEDFVSFIVLRNF 7366

QY 121 SLVGFESAFGNTRKVMRLKAAQKIFALFFPKNGCNDAICQDDSLITFSMSLDCLVG 796

DQ 737 SLVGTPLSAFGLRLPVLAEDAQRFLTALFFPEKNGCNDAICQDDSLITFSMSLDCLVG 796

797 GPREFNVTVVRNGEDSVTKQTVTFFPLDLISYRKVSTLQNQRQSRWRLACESASSTEV 856

857 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASIGNKULLKANTVSENKPRNKTPEP 916

DB 917 QLELPVKAYVMVATISHGVSTKYLNTASENTSRVMQHQQVSNLQGRSLPISLVLVPV 976

SQ	Sequence 1153 AA;	
Query Match	99.8%; Score 5862; DB 3; Length 1153;	
Best Local Similarity	99.3%; Pred. No. 0;	
Matches 1129; Conservative	7; Mismatches 1; Indels 0; Gaps 0;	
Qy	1 FNLDTENAMTFOENARGGQSVVQLQGSRRVVVGAPQEIIVANORGLSYQCDYSTGSCBPI 60	
Db	17 FNLDTENAMTFOENARGGQSVVQLQGSRRVVVGAPQEIIVANORGLSYQCDYSTGSCBPI 76	
Qy	61 RLOVPEAVNMSLGLSLAATTPCOLLAGCPVTHQTCSENTVKGCLFGLFNSLRQOPQK 120	
Db	77 RLOVPEAVNMSLGLSLAATTPCOLLAGCPVTHQTCSENTVKGCLFGLFNSLRQOPQK 136	
Qy	121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTSLFSLMOYSEEF 180	
Db	137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTSLFSLMOYSEEF 196	
Qy	181 RIHFTFKFQNNPNSLSIKPTQLLGRTHATGKRVKVVRELFNTGARKNAKILILI 240	
Db	197 RIHFTFKFQNNPNSLSIKPTQLLGRTHATGKRVKVVRELFNTGARKNAKILILI 256	
Qy	241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFRSEKSRQELNNTVASKPRDHVFQIN 300	
Db	257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFRSEKSRQELNNTVASKPRDHVFQIN 316	
Qy	301 NFEALKTIONQUREKIFAIECTQTSSESSFEHMSQEGESAITSNGPLLTSTWGSVDWAG 360	
Db	317 NFEALKTIONQUREKIFAIECTQTSSESSFEHMSQEGESAITSNGPLLTSTWGSVDWAG 376	
Qy	361 GVPLYTSKESKTFINNTRVDSMDNDAYLGVAAILLRNRVQSVILGAPRYQHIGLVAMFR 420	
Db	377 GVPLYTSKESKTFINNTRVDSMDNDAYLGVAAILLRNRVQSVILGAPRYQHIGLVAMFR 436	
Qy	421 QNTGMWESNANVKGTOIGAYFGASLCSVDVNSGSTDVLVIGAPHYEQTRGGQSVVCP 480	
Db	437 QNTGMWESNANVKGTOIGAYFGASLCSVDVNSGSTDVLVIGAPHYEQTRGGQSVVCP 496	
Qy	481 PRGORARWQCDVLYGEGQGPWRFGAALVLDGVNGDKLTDVAIGAPGEEDNRGAVILF 540	
Db	497 PRGORARWQCDVLYGEGQGPWRFGAALVLDGVNGDKLTDVAIGAPGEEDNRGAVILF 556	
Qy	541 HGTSGSGISPSHSQRIAGSKLSPRLAYFGQSLGGQDLTVGDLVDTVGAQGHVLLRSQ 600	
Db	557 HGTSGSGISPSHSQRIAGSKLSPRLAYFGQSLGGQDLTVGDLVDTVGAQGHVLLRSQ 616	
Qy	601 PVLRLVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTEDRLRQGIQSVVT 660	
Db	617 PVLRLVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTEDRLRQGIQSVVT 676	
Qy	661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLENCIEDPVSPVILRLNF 720	
Db	677 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLENCIEDPVSPVILRLNF 736	
Qy	721 SLVGTPLSAFNGRLPVLAEDAQRFTALFPPEKVCNDNTICODDLSTTFPSFMSLDCLVWG 780	
Db	737 SLVGTPLSAFNGRLPVLAEDAQRFTALFPPEKVCNDNTICODDLSTTFPSFMSLDCLVWG 796	
Qy	781 GPREFNVTVRNDEGSYRQVTFPPFLDLSYRKVSTLQNRQSORSWRLACESASSTEV 840	
Db	797 GPREFNVTVRNDEGSYRQVTFPPFLDLSYRKVSTLQNRQSORSWRLACESASSTEV 856	
Qy	841 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNTEF 900	
Db	857 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNTEF 916	
Qy	901 QLELPVKYAVVMVTSYHGVSTKYLNFTASENTRVMQHVQVSNLQSRPISLVLVVPV 960	
Db	917 QLELPVKYAVVMVTSYHGVSTKYLNFTASENTRVMQHVQVSNLQSRPISLVLVVPV 976	
Qy	961 RINQTVIWDNRQVTFPSENLSSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCORIQCDIP 1020	
Db	977 RINQTVIWDNRQVTFPSENLSSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCORIQCDIP 1036	

Qy	1021 PFGIOBEFNATLKNLSFDWYIKTSHNHLILIVSTAEILFNDSVFTLLPGQAFVRSOTET 1080	
Db	1037 PFGIOBEFNATLKNLSFDWYIKTSHNHLILIVSTAEILFNDSVFTLLPGQAFVRSOTET 1096	
Qy	1081 KVZPFFVNPPLPIVGVSSVGGLLILITAAALYKLGFFKQYKDMSEGGPPGABPQ 1137	
Db	1097 KVZPFFVNPPLPIVGVSSVGGLLILITAAALYKLGFFKQYKDMSEGGPPGABPQ 1153	
RESULT 3		
AAU80252		
ID	AAU80252 standard; protein; 1153 AA.	
AC	AAU80252;	
DT	15-JUL-2002 (first entry)	
DE	Human integrin 1 alpha-M subunit protein.	
KW	Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;	
KW	inflammatory disease; autoimmune disorder; Crohn's disease;	
KW	human immunodeficiency virus; HIV; myocardial infarction;	
KW	Sjorgen's syndrome; rheumatoid arthritis.	
OS	Homo sapiens.	
XX	Key Location/Qualifiers	
PH	Misc-difference 499..500	
FT	/note= "Encoded by GGG CAG AGG"	
FT		
XX	WO200218583-A2.	
XX	07-MAR-2002.	
XX	31-AUG-2001; 2001WO-US027227.	
PP	01-SEP-2000; 2000US-0229700P.	
PR	(BLOO-) CENT BLOOD RES INC.	
XX	Springer TA, Shimoaka M, Lu C;	
XX	WPI; 2002-382964/41.	
DR	N-PSDB; ABK50046.	
XX	Modified integrin-I or integrin I-like domain polypeptide useful as an	
PT	immunogen to produce antibodies specific to polypeptide, comprises a	
PT	disulfide bond such that polypeptide is stabilized in a desired	
PT	conformation.	
XX	Disclosure; Page 109-112; 112pp; English.	
PS		
XX	This invention relates to a modified integrin-I or integrin I-like domain	
CC	polypeptide comprising at least one disulfide bond so that the domain is	
CC	stabilized in a desired conformation. The polypeptide of the invention	
CC	may have antiinflammatory or immunosuppressive activities. The	
CC	polypeptides of the invention have an open conformation and are useful as	
CC	immunogens to produce antibodies that selectively bind to integrin I-	
CC	domain; and for identifying a modulator of integrin activity or of	
CC	interaction of an integrin and a cognate ligand. The polypeptide of the	
CC	invention, or antibodies (preferably anti-LFA-1 antibody) is useful for	
CC	treating or preventing an integrin mediated disorder which is an	
CC	inflammatory or autoimmune disorder in a subject and for inhibiting the	
CC	binding of an integrin to a cognate ligand such as Crohn's disease,	
CC	Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic	
CC	composition comprising the peptide of the invention is useful for	
CC	treating an integrin mediated disorder in a subject. The polypeptides	
CC	and/or active or antigenic fragments are useful as reagents for diagnosis	
CC	of integrin-mediated disorders. The present sequence represents the human	
CC	integrin-1 alpha-M protein subunit used to generate the mutant	
CC	polypeptides of the invention	

Query Match

Best Local Similarity 99.8%; Score 5862; DB 5; Length 1153;

Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Query Match

Best Local Similarity 99.3%; Pred. No. 0;

Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY	Sequence	1153 AA;
QY	1 FNDLTENAMTFQGNARGFGSGVVOIQGSRVVVGAPQEIIVANQRGLSQCCDYSGSCPEI 60	
DB	17 FNDLTENAMTFQGNARGFGSGVVOIQGSRVVVGAPQEIIVANQRGLSQCCDYSGSCPEI 76	
QY	61 RLQVPVEAVNMSLGLSLAATTPPQLACGPVTHTCSENTYVVKGLCFPLGSNLRQQPOK 120	
DB	77 RLQVPVEAVNMSLGLSLAATTPPQLACGPVTHTCSENTYVVKGLCFPLGSNLRQQPOK 136	
QY	121 FPALRGCGQEDSDIAFLVDGSGSIIPHDPRAKFPISTWMOIKSKTLESLMYSEEP 180	
DB	137 FPALRGCGQEDSDIAFLVDGSGSIIPHDPRAKFPISTWMOIKSKTLESLMYSEEF 196	
QY	181 RIHFTKEFQNPNPSLKIPITOLLGRTHATTGIRKVVRLEFNITNGARNNAFKILILI 240	
DB	197 RIHFTKEFQNPNPSLVKPIITOLLGRTHATTGIRKVVRLEFNITNGARNNAFKILAVI 256	
QY	241 TDEKFGDPLGHEDVIPADREGVIRYVIGGDAPFRSEKSKQELNTVASKPPRHVQFIN 300	
DB	257 TDEKFGDPLGHEDVIPADREGVIRYVIGGDAPFRSEKSKQELNTIASKPDRHVQFIN 316	
QY	301 NFEALKTIQNLREKIFALEGTQTSSSPFHESMQEGFSAAITNSGELLSTGVSYDNAG 360	
DB	317 NFEALKTIQNLREKIFALEGTQTSSSPFHESMQEGFSAAITNSGELLSTGVSYDNAG 376	
QY	361 GVLYTSKSKSTFINMTRVDSMDNAYLGAAAAIILNRVSLSVLGAPRYOHI GLVAMFR 420	
DB	377 GVLYTSKSKSTFINMTRVDSMDNAYLGAAAAIILNRVSLSVLGAPRYOHI GLVAMFR 436	
QY	421 QNTGMVESNAVKGTOIGAYFGASCSVDVDSNGSTDVLVIGAPHYYEOTRGGOVVCPL 480	
DB	437 QNTGMVESNAVKGTOIGAYFGASCSVDVDSNGSTDVLVIGAPHYYEOTRGGOVVCPL 496	
QY	481 PRGORARWCDAVLGYEQCPMGREFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540	
DB	497 PRGORARWCDAVLGYEQCPMGREFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556	
QY	541 HGTSGSGISPHSQRITAGSLSPRIYFOQSLSGGQDLTMGLVDLTVGAQHVLRLRSQ 600	
DB	557 HGTSGSGISPHSQRITAGSLSPRIYFOQSLSGGQDLTMGLVDLTVGAQHVLRLRSQ 616	
QY	601 PVLRVKAIMEFPREVARNVFECNDQVVKAGEVRVCLHVQKSTRDLREGQIQSVWT 660	
DB	617 PVLRVKAIMEFPREVARNVFECNDQVVKAGEVRVCLHVQKSTRDLREGQIQSVWT 676	
QY	661 YDLALDSGRPHSAVFNETKNSTRQTQVLGTTCETKLKLPNCIEDPSPVILRLNF 720	
DB	677 YDLALDSGRPHSAVFNETKNSTRQTQVLGTTCETKLKLPNCIEDPSPVILRLNF 736	
QY	721 SLVGTPLSAFGLNRPVLAEDAQLFETALPFPEKNGCNINI CODDLSTIFFSMJLDCLVWG 780	
DB	737 SLVGTPLSAFGLNRPVLAEDAQLFETALPFPEKNGCNINI CODDLSTIFFSMJLDCLVWG 796	
QY	781 GPREFNVTVTRNDGSDSYRTQVTFPPFDLDSYRKVSTLQNRQSRLWELACESASTEV 840	
DB	797 GPREFNVTVTRNDGSDSYRTQVTFPPFDLDSYRKVSTLQNRQSRLWELACESASTEV 856	
QY	841 SGALKSTCSINHIIPPENSEVTNITFDVDSKASLGKLLKANVTSENMPNTKTEP 900	
DB	857 SGALKSTCSINHIIPPENSEVTNITFDVDSKASLGKLLKANVTSENMPNTKTEP 916	
QY	901 QLELPVKIAYVMVTSKHGYSKYLNFTASENTRVWQHYOVSNLQORSLPISLVFLVPV 960	
DB	917 QLELPVKIAYVMVTSKHGYSKYLNFTASENTRVWQHYOVSNLQORSLPISLVFLVPV 976	
QY	961 RLMQTVIMWRPQVTFSENISSCTCHKERLPSSHDFLABLRKAPVNNVSIACVQRIQCIDIP 1020	
DB	977 RLMQTVIMWRPQVTFSENISSCTCHKERLPSSHDFLABLRKAPVNNVSIACVQRIQCIDIP 1036	

QY	Sequence	1153 AA;
QY	1021 PFQIOEEFNATLKGMLSFDMWIKTSHNELLIIVSTABILFNDSVTLPLPGQGAFVRSOTET 1086	
DB	1037 PFQIOEEFNATLKGMLSFDMWIKTSHNELLIIVSTABILFNDSVTLPLPGQGAFVRSOTET 1096	
QY	1081 KYEPFEVNPNPLIIVGSSVGGILLIALITALYKLGFFPKROYKDMSEGPPGAEPO 1137	
DB	1097 KYEPFEVNPNPLIIVGSSVGGILLIALITALYKLGFFPKROYKDMSEGPPGAEPO 1153	

RESULT 5

AA014428

ID AA014428 standard; protein; 1153 AA.

XX AAO14428;

AC AAO14428;

XX

DT 03-MAY-2002 (first entry)

XX

Integrin Mac-1 alpha subunit.

DE

Mac-1; integrin alpha subunit; variant integrin inserted domain protein; open conformation; integrin related inflammatory disorder;

KW

integrin related immunological disorder; rheumatoid arthritis; ischaemia; reperfusion; hypovolemic shock; infarction; cerebral shock;

KW

viral infection; cancer; gene therapy; vaccine;

KW

bioactive agent screening.

XX

Unidentified.

CS

XX

WO200204521-A2.

PN

17-JAN-2002.

PD

XX

09-JUL-2001; 2001WO-US021805.

PF

XX

07-JUL-2000; 2000US-0216600P.

PR

XX

(CALY) CALIFORNIA INST OF TECHNOLOGY.

PA

(BLOO-) CENT BLOOD RES.

PA

Springer T;

PI

XX

WPI; 2002-148167/19.

DR

XX

New integrin I domain protein having alteration in at least 2 noncontiguous regions and exits in an open conformation, useful for treating, preventing or suppressing inflammatory or immunological disorders.

PT

XX

Example 1; Fig 1P; 90pp; English.

PS

XX

The invention comprises structurally biased variant integrin inserted (I) domain proteins, wherein the alterations to the protein occur in at least two noncontinuous regions. Specifically the variant integrin I domain proteins are structurally biased to exist in the open conformation, thereby altering the binding ability of the protein. The invention also comprises nucleic acids encoding the variant integrin I domain proteins. The integrin I domain proteins and nucleic acids are useful for treating, preventing or suppressing integrin related inflammatory and immunological disorders (e.g. rheumatoid arthritis). The variant integrin I domain proteins and nucleic acids can also be used for treating: ischaemia/reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral infection; and cancer. The variant integrin I domain nucleic acids and proteins may be used in gene therapy, as vaccines and to screen for bioactive agents. The present amino acid sequence represents the Mac-1 alpha subunit of integrin

CC

Sequence 1153 AA;

XX

Query Match 99.8%; Score 5862; DB 5; Length 1153;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARFGQSVVOLQSGRRVVVGAPOEIVAAAMORGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARFGQSVVOLQSGRRVVVGAPOEIVAAAMORGSLYQCDYSTGSCPEI 76
QY 61 RLOVPVAVNMSLGLSLAATTSPOILLACGPTVHOTCSNTYVKGICPLFGSNLRQPOOK 120
DB 77 RLOVPVAVNMSLGLSLAATTSPOILLACGPTVHOTCSNTYVKGICPLFGSNLRQPOOK 136
QY 121 FPEALRGCPQEDSDIAFLDGGSGSIIPHPFRRAKBEISTVMEQLKSKSLFSLMQYSEEP 180
DB 137 FPEALRGCPQEDSDIAFLDGGSGSIIPHPFRRAKBEISTVMEQLKSKSLFSLMQYSEEP 196
QY 181 RIHFTPKBFQNNPNPSLKIPIQLLGRTHATGIRKVVRELPNIINGARKNAFKILLI 240
DB 197 RIHFTPKBFQNNPNPSLKIPIQLLGRTHATGIRKVVRELPNIINGARKNAFKILLI 256
QY 241 TDEKFGDPLGYEDVPEADREGVIRVIGVGDFAFRSEKSRQELNIVASKPPRDHVFQIN 300
DB 257 TDEKFGDPLGYEDVPEADREGVIRVIGVGDFAFRSEKSRQELNIVASKPPRDHVFQIN 316
QY 301 NFALKTIONOLREKIPAIISGTOTGSSSSPEHEMSQSGPSAALTNGPILLSVGSYDWAG 360
DB 317 NFALKTIONOLREKIPAIISGTOTGSSSSPEHEMSQSGPSAALTNGPILLSVGSYDWAG 376
QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGVAAILLRNRVQSLVIGAPRYOHLGLVAMER 420
DB 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGVAAILLRNRVQSLVIGAPRYOHLGLVAMER 436
QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGRGQSVCP 480
DB 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGRGQSVCP 496
QY 481 PRGORARWQDAVLVYGQGGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNKGAVILF 540
DB 497 PRGORARWQDAVLVYGQGGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNKGAVILF 556
QY 541 HGTSGSGISPSHQSRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHQSRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLVKAIEMFNPREVARNVPCNDQVVKGEAGEVAVCLHVQKSTDRLRREGOISVVT 660
DB 617 PVLVKAIEMFNPREVARNVPCNDQVVKGEAGEVAVCLHVQKSTDRLRREGOISVVT 676
QY 661 YDLALDSGRPHRAVFNENKSTTRQTVLGLTQTCETKLQLPNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPHRAVFNENKSTTRQTVLGLTQTCETKLQLPNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAPGNLSPVLAEDQRLFTALPPEKKNCGNDNICQDDLSITPFSMSLDCLVVG 780
DB 737 SLVGTPLSAPGNLSPVLAEDQRLFTALPPEKKNCGNDNICQDDLSITPFSMSLDCLVVG 796
QY 781 GPREFNVTVVRNDEGDSYRQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDEGDSYRQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSENNMPTNKTEP 900
DB 857 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSENNMPTNKTEP 916
QY 901 QLELPVKYAVVMVTSHGVSFKYLNFTASENTSRVMOHYQVSNLQORSIPISLVFLVPV 960
DB 917 QLELPVKYAVVMVTSHGVSFKYLNFTASENTSRVMOHYQVSNLQORSIPISLVFLVPV 976
QY 961 RLNQTIVWDRPQVTFSENLSSTCHTKERLPKSHSDPFLAELRKAPVNVCSIAVCORIQDIP 1020
DB 977 RLNQTIVWDRPQVTFSENLSSTCHTKERLPKSHSDPFLAELRKAPVNVCSIAVCORIQDIP 1036
QY 1021 PFGIOEEFNATLKGNSLFDWIKTSHNHLIVSTABILLFNDSVFTLLPGQAFVRSOTET 1080
DB 1037 PFGIOEEFNATLKGNSLFDWIKTSHNHLIVSTABILLFNDSVFTLLPGQAFVRSOTET 1096
QY 1081 KVEPPEVFNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMWSEGGPPGAEPQ 1137

DB 1097 KVEPPEVFNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMWSEGGPPGAEPQ 1153
RESULT 6
ADD25615
ID ADD25615 standard; protein; 1153 AA.
XX ADD25615;
AC ADD25615;
XX
DT 15-JAN-2004 (first entry)
XX Binding domain-immunoglobulin fusion protein-associated protein #85.
DE Binding domain; immunoglobulin; fusion protein; cytostatic;
KW antirheumatic; immunosuppressive; antidiabetic; antithyroid;
KW neuroprotective; hinge region; immunoglobulin heavy chain;
KW CH2 constant region; CH3 constant region; IgG1;
KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
KW malignant condition; B-cell disorder; melanoma; sarcoma;
KW rheumatoid arthritis; myasthenia gravis; Grave's disease;
KW type 1 diabetes mellitus; multiple sclerosis; autoimmune disease.
XX Unidentified.
OS
XX US2003118592-A1.
XX
PD 26-JUN-2003.
XX
XX 25-JUL-2002; 2002US-00207655.
XX
XX 17-JAN-2001; 2001US-0367358P.
PR 17-JAN-2002; 2002US-00053530.
PR 03-JUN-2002; 2002US-0385691P.
XX
XX (GENE-) GENE-CRAFT INC.
XX
XX Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
PI 2003-801317/75.
XX
XX New binding domain-immunoglobulin fusion protein, useful for treating a
subject having or suspected of having a malignant condition or a B-cell
disorder, e.g. melanoma, Grave's disease or autoimmune disease.
XX
XX Disclosure; SEQ ID NO 176; 157pp; English.
XX
XX The invention relates to a binding domain-immunoglobulin fusion protein
comprising a binding domain polypeptide that is fused to an
immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain
CH2 constant region polypeptide that is fused to the hinge region
polypeptide, and an immunoglobulin heavy chain CH3 constant region
polypeptide that is fused to the CH2 constant region polypeptide. The
hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin hinge
region polypeptide; a mutated human IgG1 immunoglobulin hinge
region polypeptide, derived from (a) having 3 or more cysteine residues;
where the mutated human IgG1 immunoglobulin hinge region polypeptide
contains 2 cysteine residues, where the first cysteine is not mutated; a
mutated human IgG1 immunoglobulin hinge region polypeptide, derived from
(a) having 3 or more cysteine residues, where the mutated human IgG1
immunoglobulin hinge region polypeptide contains no more than one
cysteine residue; and a mutated human IgG1 immunoglobulin hinge region
polypeptide, derived from (a) having 3 or more cysteine residues; where
the mutated human IgG1 immunoglobulin hinge region polypeptide contains
no cysteine residues. The binding domain-immunoglobulin fusion protein is
capable of at least one immunological activity comprising antibody
dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The
binding domain polypeptide is capable of specifically binding to an
antigen. Also included are an isolated polynucleotide encoding the
binding domain-immunoglobulin fusion protein, a recombinant expression
construct comprising the polynucleotide (operably linked to a promoter),
a host cell transformed or transfected with a recombinant expression
construct, producing the binding domain-immunoglobulin fusion protein, a

pharmaceutical composition comprising the binding domain-immunoglobulin fusion protein or polynucleotide and a carrier, and treating a subject having or suspected of having a malignant condition or a B-cell disorder. The binding domain-immunoglobulin fusion protein is useful for treating a subject having or suspected of having a malignant condition or a B-cell disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis, myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple sclerosis or autoimmune disease. The present sequence is a binding domain -immunoglobulin fusion protein-associated protein sequence. Note: The sequence data for this patent formed part of the printed specification and is also available in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not identified the sequences in the printed specification by their SEQ ID number therefore none of the sequences can be explicitly identified.

Sequence 1153 AA;

Query Match 99.8%; Score 5862; DB 7; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

1 FNLDTENAMTFOENARGFGQSVVQLGGRRVVGAPQEIWAANQKSLYQCDYSTGSCPEI 60
17 FNLDTENAMTFOENARGFGQSVVQLGGRRVVGAPQEIWAANQKSLYQCDYSTGSCPEI 76
61 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHOTCSNTYVKGCLTFLGSLNRQQPQK 120
77 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHOTCSNTYVKGCLTFLGSLNRQQPQK 136
121 FPEALRGCPQSDIAFLVDGSGSIIPHDFRAKEFISTVMEQLKSKTLFSLMYSEEF 180
137 FPEALRGCPQSDIAFLVDGSGSIIPHDFRANKFVSTVWEQLKSKTLFSLMYSEEF 196
181 RIHFTFKEFQNNPNRSLKPIITQLGRTHATGIRKVVRELFTNGARKNAFKILILI 240
197 RIHFTFKEFQNNPNRSLKPIITQLGRTHATGIRKVVRELFTNGARKNAFKILIVI 256
241 TIGEXFGDPLGYEDVTPADRGVIRYVIGVGDAFRSEKSRQELMTVASKPRDHFVQIN 300
257 TIGEXFGDPLGYEDVTPADRGVIRYVIGVGDAFRSEKSRQELMTVASKPRDHFVQIN 316
301 NFEALKTIQNLREKIPALEGTQTSSESFHEMSQEGFSAAITNGPLSTVGSYDWAG 360
317 NFEALKTIQNLREKIPALEGTQTSSESFHEMSQEGFSAAITNGPLSTVGSYDWAG 376
361 GVFLYTSKSKSTFINNTRVDSMDNAYLGAAAAIILRNVRQSLVIGAPRYQHIGLVAMPR 420
377 GVFLYTSKSKSTFINNTRVDSMDNAYLGAAAAIILRNVRQSLVIGAPRYQHIGLVAMPR 436
421 QNTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVILGAPHYEQTGQGVSVCP 480
437 QNTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVILGAPHYEQTGQGVSVCP 496
481 PRQQRARWQCDVLYGEOQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
497 PRQQRARWQCDVLYGEOQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTGCAQGHVLLRSQ 600
557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTGCAQGHVLLRSQ 616
601 PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLRVOKSTRDLRLEGQIQSVYT 660
617 PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLRVOKSTRDLRLEGQIQSVYT 676
661 YDLALDSGPHSRVAVNETKSTRQTVLGTOTCETLKLQPNCTEDPVPVILRLNF 720
677 YDLALDSGPHSRVAVNETKSTRQTVLGTOTCETLKLQPNCTEDPVPVILRLNF 736
721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALPPFEKNCNDNICQDDLSTIFFSFMSLDCLVVG 780
737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALPPFEKNCNDNICQDDLSTIFFSFMSLDCLVVG 796

QY 781 GPREFNVTVTRMDGEDSYRTQVTFPPFLDLSYRKVSTLQNSQRSWRLACSSASSTEV 840
Db 797 GPREFNVTVTRMDGEDSYRTQVTFPPFLDLSYRKVSTLQNSQRSWRLACSSASSTEV 856
QY 841 SGALKSTSCSINHPFIPENSEVTNITFDVDSKASLGNKLLKANTSENKMPSTNKTET 900
Db 857 SGALKSTSCSINHPFIPENSEVTNITFDVDSKASLGNKLLKANTSENKMPSTNKTET 916
QY 901 QLELPVAVYVAVVMTSHGVSTKYLNFNTASNTSRVMOHQYQVSNLQORSLSPLSVFLVPV 960
Db 917 QLELPVAVYVAVVMTSHGVSTKYLNFNTASNTSRVMOHQYQVSNLQORSLSPLSVFLVPV 976
QY 961 RLNQTIVDRPQVTPSENLSTCHTKERLPSHSDFLAELKAPVNCISIAVCQRIQCDIP 1020
Db 977 RLNQTIVDRPQVTPSENLSTCHTKERLPSHSDFLAELKAPVNCISIAVCQRIQCDIP 1036
QY 1021 PFGQIESFNATLKGNLISDFMDWIKTSHNHLIIVSTABILFNDVSFTLLPGQGAFFVRSOTET 1080
Db 1037 PFGQIESFNATLKGNLISDFMDWIKTSHNHLIIVSTABILFNDVSFTLLPGQGAFFVRSOTET 1096
QY 1081 KVPEFVNPPLPLIVGSSVGGLLALLITAAALYKLGFFKQYKDMSEGGPPGAEPO 1137
Db 1097 KVPEFVNPPLPLIVGSSVGGLLALLITAAALYKLGFFKQYKDMSEGGPPGAEPO 1153

RESULT 7
AAR04136
ID AAR04136 standard; protein; 1153 AA.
XX AAR04136;
AC AAR04136;
XX AAR04136;
DT 25-MAR-2003 (revised)
DT 07-SEP-1990 (first entry)
DE Alpha subunit of Mac-1 leukocyte adhesion receptor.
XX Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
XX non-specific defence system; integrin gene superfamily.
OS Synthetic.
XX
Key Location/Qualifiers
Region 1..16
FT /label= signal peptide
FT Modified-site 86..88
FT /label= putative N-glycosylation site
FT Modified-site 240..242
FT /label= putative N-glycosylation site
FT Modified-site 391..393
FT /label= putative N-glycosylation site
FT Modified-site 469..471
FT /label= putative N-glycosylation site
FT Modified-site 693..695
FT /label= putative N-glycosylation site
FT Modified-site 697..699
FT /label= putative N-glycosylation site
FT Modified-site 735..737
FT /label= putative N-glycosylation site
FT Modified-site 802..804
FT /label= putative N-glycosylation site
FT Modified-site 881..883
FT /label= putative N-glycosylation site
FT Modified-site 901..903
FT /label= putative N-glycosylation site
FT Modified-site 912..914
FT /label= putative N-glycosylation site
FT Modified-site 941..943
FT /label= putative N-glycosylation site
FT Modified-site 947..949
FT /label= putative N-glycosylation site
FT Modified-site 979..981
FT /label= putative N-glycosylation site
FT Modified-site 994..996

FT Modified-site /label= putative N-glycosylation site
 FT 1022..1024
 FT /label= putative N-glycosylation site
 FT 1045..1047
 FT Modified-site /label= putative N-glycosylation site
 FT 1051..1053
 FT Modified-site /label= putative N-glycosylation site
 FT 1076..1078
 FT Modified-site /label= putative N-glycosylation site
 FT 1106..1134
 FT Region /label= putative transmembrane region
 XX EP364690-A.
 XX
 XX 25-APR-1990.
 XX
 XX 17-AUG-1989; 89EP-00115159.
 XX
 XX 23-AUG-1988; 88US-00235353.
 PR
 PR 09-MAR-1989; 89US-00321239.
 XX
 XX (DAND) DANA FARBER CANCER INST INC.
 PA
 PI Springer TA, Corbi A;
 DR WPI: 1990-125938/17.
 DR N-PSDB; AAQ04043.
 XX
 XX New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating
 FT inflammation and viral infections, and in diagnosis.
 FT
 PS Disclosure; Page 2; -pp; English.
 XX
 CC Mac-1 alpha subunit is involved in the response to inflammation, i.e.
 CC recognition of and migration to sites of inflammation. It also attaches
 CC to cellular substrates as part of this function making it useful in
 CC visualising endothelial tissue. Mac-1 is a member of the Integrin Gene
 CC superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25
 CC -MAR-2003 to correct PA field.)
 XX
 SQ Sequence 1153 AA;
 Query Match 99.6%; Score 5852; DB 2; Length 1153;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ENLDTENAMTFQENARGFGQSVVQLQGSRRVWVGAPQEIIVAAHQSGLYQCDYSTGSCPEI 60
 DB 17 ENLDTENAMTFQENARGFGQSVVQLQGSRRVWVGAPQEIIVAAHQSGLYQCDYSTGSCPEI 76
 QY 61 RLOVPVEANVMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 120
 DB 77 RLOVPVEANVMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 136
 QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFIPTVMEOLKESKTLFSLMOYSEEP 180
 DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFIPTVMEOLKESKTLFSLMOYSEEP 196
 QY 181 RIHFTKEFQNNPRSLKPIITQLGRTHATGIRKVVRELFTNITNGARKNAFKILILI 240
 DB 197 RIHFTKEFQNNPRSLKPIITQLGRTHATGIRKVVRELFTNITNGARKNAFKILIVI 256
 QY 241 TDGCKGDPGLGYEDVTPRADREGVIRVIGVGAPRSEKSRQELNTVASKPRDHDVQFN 300
 DB 257 TDGCKGDPGLGYEDVTPRADREGVIRVIGVGAPRSEKSRQELNTVASKPRDHDVQFN 316
 QY 301 NFALKTIQNLREKIFAIBGTQTGSSSPFHEMSQEGFSAATISNGPLLTSTVGSYDNAG 360
 DB 317 NFALKTIQNLREKIFAIBGTQTGSSSPFHEMSQEGFSAATISNGPLLTSTVGSYDNAG 376
 QY 361 GVFLYTSKEKSTINTRVDSNDNDAYLVGAAAIILIRNVQSLVLCAPRYOHIGLVAMFR 420
 DB 377 GVFLYTSKEKSTINTRVDSNDNDAYLVGAAAIILIRNVQSLVLCAPRYOHIGLVAMFR 436

RESULT 8

AAR07120

ID AAR07120 standard; protein; 1163 AA.

XX AC AAR07120;

XX DT 25-MAR-2003 (revised)

XX DT 05-FEB-1991 (first entry)

XX DE p150.95 alpha subunit encoded by clone lambdaX47.

XX KW p150,95 leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia;
rhinovirus.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Region 1..19

XX FT /label= signal peptide

XX FT Region 20..44

XX FT /label= N-terminus

XX FT Modified-site 61..63

XX FT /label= glycosylation site

FT Modified-site 89..91
 FT /label= glycosylation site
 FT Modified-site 385..387
 FT /label= glycosylation site
 FT Modified-site 392..394
 FT /label= glycosylation site
 FT Modified-site 697..699
 FT /label= glycosylation site
 FT Modified-site 735..737
 FT /label= glycosylation site
 FT Modified-site 899..901
 FT /label= glycosylation site
 FT Modified-site 904..906
 FT /label= glycosylation site
 FT Modified-site 939..941
 FT /label= glycosylation site
 FT Modified-site 1050..1052
 FT /label= glycosylation site
 FT Domain 1108..1133
 FT /label= transmembrane
 FT XX
 PN W09010646-A.
 XX
 XX 20-SEP-1990.
 XX
 XX 09-MAR-1990; 90W0-US001257.
 XX
 XX 09-MAR-1990; 90W0-US001257.
 XX
 XX (DAND) DANA FARMER CANCER INST INC.
 XX PA
 XX Corbi AA, Springer TA;
 XX
 XX WPI; 1990-304985/40.
 DR N-PSDB; AAQ06068.
 XX
 CC Treatment of viral esp. rhino-viral infection - by admin. of alpha sub-
 PT unit of p150.95 cell surface adhesion receptor, opt. together with a beta
 PT chain of CD-18 family.
 PT
 XX
 XX Disclosure; Fig 3; 59pp; English.
 XX
 CC Clone lambda x47 was isolated from a cDNA library constructed from total
 CC RNA extracted from phorbol myristate acetate stimulated HL-60
 CC myelomonocytic cells. The library was screened with oligonucleotide
 CC probes based on tryptic peptide fragments of p150.95. The sequence can be
 CC attached to appropriate control elements and expressed in prokaryotic and
 CC eukaryotic cells. The protein can be used to treat or prevent rhinoviral
 CC infection because it interacts with ICAM-1 and inhibits cell-virus
 CC attachment. It can also be used as an anti-inflammatory agent. See also
 CC AAQ06063-4, AAQ06068, AAQ07121-8 and AAQ07152-6. (Updated on 25-MAR-2003
 CC to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 1163 AA;
 Query Match 58.9%; Score 3460; DB 2; Length 1163;
 Best Local Similarity 60.9%; Pred. No. 5.4e-280;
 Matches 687; Conservative 141; Mismatches 295; Indels 6; Gaps 4;
 QY 1 FNLDENAMTQENARFGQSVVLOGSRVVGAPQIVAAVNRGSLGYOCDYSTGSCPEI 60
 DB 20 FNLDTEILTAFRVDGAGDSVVGAPQIVAAVNRGSLGYOCDYSTGSCPEI 79
 QY 61 RLQVPEAVNMSLGLAATTPSPQILACGPTVHOTCSENYVKGCLFLGSLNRLOQPOK 120
 DB 80 GLQVPEAVNMSLGLAATTPSPQILACGPTVHOTCSENYVKGCLFLGSLNRLOQPOK 137
 QY 121 FPEALRGQEDSDIAFLVDSGSIIPHDPRKAFPISTVMEQLKSKTFLSLMVSSEEP 180
 DB 138 LPVSRQECPEQDQIVFLIDGSGISRRNFATMNFVRAVISQFSPSTQFSLMQPSNKF 197
 QY 181 RIHFTFEKFNPNPRSLIKPITOLLGRTHATGIRKRVRELFNITNGARKNAFKILLI 240

Db 198 QTHPTPEFRFRTSNPLSLLASVHQLQGYTATATQNVVHRLPHASYGAREDAKILIVI 257
 QY 241 TDGEKFGDPIGYEDVPEADREGRVIRYVGVGDAPRSEKROELNVTASPPRKHVPQIN 300
 Db 258 TDGKEGSDLDYKDVIPWADAAGIIRYAIWGLAFQNRNSWKEINLDTASPSQEHIPKVE 317
 QY 301 NFEALKTIQNLREKIFAIECTGTGSSSPFEHMSQSGFSAALITNSGPLLSTVGSVDWAG 360
 Db 318 DFDALKIQLQNLKEKIFAIECTGTGSSSPFEHMSQSGFSAALITNSGPLLSTVGSVDWAG 377
 QY 361 GVFLYTSKSKSTFPIWTRVDSMDNDAYLGYAAAILLRNVOSLIVGAPRYOHLVAMPR 420
 Db 378 GAFLYPPNMSPTFINNGQENVMDRDSYLGYSTELALWKGQSLVGLAPRYOHTKAVIFT 437
 QY 421 QNTGMWESNANVKTQICAYFGASLCSVDVDSNGSTDVLIGAPHYEYQNRGGQSVCP 480
 Db 438 QVSRQWKAETVGTQIGSYFGASLCSVDVDSNGSTDVLIGAPHYEYQNRGGQSVCP 497
 QY 481 PRQARWOCDAVLYGEGQCPWGRFGAALTVDVNGDKLTDVAIGPCEDEIRGAYILF 540
 Db 498 PRGWR-RWMCDAVLYGEGQCPWGRFGAALTVDVNGDKLTDVWVGAPGSEENRGAAYLF 556
 QY 541 HGTSGGISTPSHSORIASGLSPRLQYFGOSLGGQDLTMDGLVDTLVGAQGHVLLRSQ 600
 Db 557 HGVLGPSISPSHSORIASGLSPRLQYFGQALSGGQDLTODGLVLAAGARGQVLLLR 616
 QY 601 PVLRYKAIMENPREVARNVPECNDQVWKGKEAGEVURVCLHVKSTDRLEKQIQSVVT 660
 Db 617 PVLWVGVSQWQIFABIPRSAFECEQVVSQTLVQSNICLYIDKRSKNLGLSGRDLSSVT 676
 QY 661 YDLALDSCRPHSRVAFNETKNSRRTQVTLGLTQTCETLKLQLPNCIEDPVSRTVLRNF 720
 Db 677 LDLDALDQGLSPRATFQETKNSRSLRVRLGLAKHCENFLLLPSCVEDSVTITRLNF 736
 QY 721 SLVGTPLSAPGNLPAEDAOQLFTALPFPEKNCNDNTCQDDLSITFSFMSLDCIIVG 780
 Db 737 TLVSKPILAFNLSPMLAALAQRYFTASLPFEKNCADHICQDNLGSGFSFGKLSLVG 796
 QY 781 GPREFNVTVTVRNDEGDSYRTQVTFPPDLDSYRKVSTLQNRQSQRSWRLACESASSTEV 840
 Db 797 SNLEINAEVWVNDGDSYGTITTFSHPAGLSYRYVAEGQKQSLSLHTCDSA--PVA 854
 QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSASLGNKLLKANYTSENMRNTKTEF 900
 Db 855 SQGTWSTSCRNHLIFRGAQITFLATFDVSPKAVLGDRILLTANVSSNNPTSKITF 914
 QY 901 QLELPVYAVVWVTVSHGVSTKYLNPITAS-ENTSRVMOHOYQVSNLQSRSLPSLPLVP 959
 Db 915 QLELPVYAVVTVVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQDRDLPVSNFWVP 974
 QY 960 VRLNQTVINRDPQVTFSENLSSTCHTKERLPSHSDFLAEKAPVNCSTAVQRIQCDI 1019
 Db 975 VELNQAVWMDVSVHPQNPSPSLRCSQKIAPPASDDEFLAHIQKWPVLDSCSAGCLRPRCDV 1034
 QY 1020 PFFGIQBEFNATLKNLSFDWYIKTSHNHLIIVSTABILFNDSVFTLLPQCGAFVRSQTE 1079
 Db 1035 PSFSVQSELDFTLKNLSFGWVRQILQKKVSVVVAEITEDTSVYSLRQQRERFMAQIT 1094
 QY 1080 TKVFPPEVPEPLPIVGVSSVGGILLIALLIITAAALKYGLFFKROKDMWSE 1128
 Db 1095 TVLEKYKVHNPPLIVGSSIGALLIALLIITAVLYKGVFFKROKEMWEE 1143
 RESULT 9
 ID AAW65091
 XX AAW65091 standard; protein; 1163 AA.
 AC AAW65091;
 XX
 XX 28-SEP-1998 (first entry)
 XX Human Beta-integrin CD11c subunit protein.
 DE
 XX

KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW type-1 diabetes; atherosclerosis; multiple sclerosis; asthma;
KW lung inflammation; acute respiratory distress syndrome; CD11c subunit;
KW rheumatoid arthritis.
OS Homo sapiens.
XX US5728533-A.
XX 17-MAR-1998.
XX 07-JUN-1995; 95US-00485618.
XX 23-DEC-1993; 93US-00173497.
XX 05-AUG-1994; 94US-00286889.
XX 21-DEC-1994; 94US-00362652.
XX (ICOS-) ICOS CORP.
XX Van Der Vieren M, Gallatin W;
XX WPI; 1998-206565/18.
XX Screening assay for modulators of integrin binding - using immobilised or
XX labelled alpha-d polypeptide, useful for, e.g. treating type-1 diabetes.
XX Example 5; Fig 1A-D; 106pp; English.
XX This sequence represents a human beta-integrin CD11c subunit which is
XX used to describe a method for identifying compounds that modulate the
XX interaction of the beta-integrin alpha-d subunit with a binding partner
XX of alpha-d which involves contacting an alpha-d polypeptide with an alpha
XX -d binding partner, one of which is immobilised and the other of which is
XX labelled, in the presence of a test compound, and determining if the
XX compound affects binding between the alpha-d polypeptide and alpha-d
XX binding partner, where the alpha-d polypeptide is alpha-d or its fragment
XX comprising the cytoplasmic, transmembrane or extracellular domain of
XX alpha-d. Compounds that modulate alpha-d binding could be used to treat
XX diseases such as type-1 diabetes, atherosclerosis, multiple sclerosis,
XX asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
XX and rheumatoid arthritis
XX Sequence 1163 AA;

Query Match 58.6%; Score 3446; DB 2; Length 1163;
Best Local Similarity 60.8%; Pred. No. 8e-279;
Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;
QY 1 FNLDENAMTFOENARGGQSVQOLQGSVVVVGAPQEIIVANQSGSYQCDYSTGSCPEI 60
DB 20 FNLDTEELTAPRVDAGFGDSVQVYANSWVVGAPQEIIVANQSGSYQCDYSTGSCPEI 79
QY 61 RLQVPVAVNMVSLGLSLAATTPQLLACGPTVHTQCSNTYKGLCLFLPGSNLRQOPQK 120
DB 80 GLQVPPVAVNMVSLGLSLAATTPQLLACGPTVHTQCSNTYKGLCLFLPGSNLRQOPQK 137
QY 121 PPEALRGCPQEDSDIAFLVQSGSIIIPDPERAKEFIPTWQELKSKTLFSLMOYSEEP 180
DB 138 LPVRQBCFQEQDIFVLIDSGSSISSENFMATMNFVRAVISQFQESTQFSLMQFNKP 197
QY 181 RIHFTFEFQNNPRLSLIKPITQLLGRTHATGIRKVVRELFTNGARKNAPKILILI 240
DB 198 QTHFTFEFRTSNFLSLASVHQLQGYTATATQVWVHLPFASGARDAIKILIVI 257
QY 241 TDGKFGDPLGYEDVPEADREGVIRVYVGDFAFRESKQELNTYASKPPRDHVFQIN 300
DB 258 TDGKKGSDLDYKQVWPADAAGIIRVYVGLAFQNNRNSWKELNDIASKPSQEHKPV 317
QY 301 NFEALKTIQNLREKIPAEIGTQGTSSSPHEMSQEGFSAATNSGPLLSTVGSYDWAG 360
DB 318 DFDALDKDQNLQKELIPAEIGTETISSSSFLEWAQEGFSAVFTPDGVLGAVCSFTMSG 377
QY 361 GVFLYTSKEKSTFFINMTVDSMDNDAYLGVAAAAILNRVQSLVGLGAPRYQHIGLVAMPR 420

DB 378 GAFLYPPNNSPTFFINMSQENVMDRDSYLGYSTELALMKGVQSLVGLGAPRYQHIGKAVIFI 437
QY 421 QNTGHWENANVKGTOIGAYFGASLCSVDVDSNSTDLVILGAPHYETGTRGGQSVVCPCL 480
DB 438 QVSRQWRMAEVIQTIGSYFGASLCSVDVDTGSDTLVILGAPHYETGTRGGQSVVCPCL 497
QY 481 PRGQARWQCDVLYGEGQPGWRFGAALTVLGDVNGDKLTDVAIGAPGEENRGAATL 540
DB 498 PRGWR-RWVCDVLYGEGQPGWRFGAALTVLGDVNGDKLTDVWIGAPGEENRGAATL 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVLDLTVGAQGHVILLRSQ 600
DB 557 HGVLGFSISPSHSQRIAGSKLSSRLQYFGQALSGQDLTQDGLVLDVLAGARGVILLRTR 616
QY 601 PVLRYKATMBEPREVARNVPCNDQVYKGEAGVRVCLHVOKSTRDLREGQIQSVVT 660
DB 617 PVLWGVSKQFIPAEIPRPAFECRQVYSEQTLVQSNICLYIDKSKNLGSRDLQSSVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRQTOVLGTQTCETLKLQLPNCIEDPVPVILRLNF 720
DB 677 LOLALAPGRLSPRATFQETKNSLSRVRLGAKACENFNLLPSCVEDSVIPIILRLNF 736
QY 721 SLVGTPLSAPGNLRLVLAEDAQRLFTALPPPEKNCNDNICODLSITFSFMSLQCLVVG 780
DB 737 TLVGPXLLAFRLRLPMLAALAOYFTASLPPEKNCADHICQDNLGISEFSFPLKSLVVG 796
QY 781 GPREFNVTVTVENDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRSMWPLACESASSTEV 840
DB 797 SNLELNAEYVWVNDGEDSYGTTITFSHPAGLSYRVVAEQKQQLRSLHLTC--CSAPV 854
QY 841 SGALKSTSCINHPFPENSEVTFNITPDVDSKSLGNKLLKANVTSENMPTNKTEF 900
DB 855 SQTWTSTSCRIHILFRGGAQTFLATPDVSPRAVLDRLLLIANVSSNNIPRTSKTIP 914
QY 901 QLELPKYAVYVWVTVSHGVSTKYLNFTAS-ENTSRVMOHQYQVSNLQGRSLPISLVFLVP 959
DB 915 QLELPKYAVYVWVSHGVSTKYLNFTAS-ENTSRVMOHQYQVSNLQGRSLPISLVFLVP 974
QY 960 VRLNQTVIWDROVTFSESNLSTCHKEPLPSHDFELARLKPAPVNCISVACQICQDI 1019
DB 975 VELNQEAVWVDEVSHPQNPSEKTAIPASDPLAHIQNPVLDSCISAGCLRRCDDV 1034
QY 1020 PFGIOEBFNATLKGNLSPDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFAVSQTE 1079
DB 1035 PFSVQGEELDFTLKGNLSEGWVRIILQKXSVSVVAEIIIFTDSVYSQLPQBAFMAQTI 1094
QY 1080 TKVPEPPEVNPPLPIVSGSVGLLLALITALYALYKLGFFKQYKMMSE 1128
DB 1095 TVLEKYKVNPIPLIVGSSIGGLLILALITAVLYKVGFFKQYKEMMBE 1143
RESULT 10
AAB07361
ID AAB07361 standard; protein; 1163 AA.
XX AAB07361;
XX 17-JAN-2001 (first entry)
XX Human CD11c protein sequence.
XX Human; macrophage infiltration inhibition; alpha d integrin;
XX leukocyte integrin; Leu-CAM; leukointegrin; immune response;
XX inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
XX atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
XX lung inflammation; acute respiratory distress syndrome; Crohn's disease;
XX rheumatoid arthritis; central nervous system injury; CD11c.
XX Homo sapiens.
XX W0200029446-A1.
XX PN

25-MAY-2000.
 16-NOV-1999; 99WO-US027139.
 16-NOV-1998; 98US-00193043.
 08-JUL-1999; 99US-00350259.
 (ICOS-) ICOS CORP.
 Gallatin M, Van Der Vieren M;
 WPI; 2000-387751/33.
 Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
 macrophage infiltration and reduce inflammation at central nervous system
 injury sites.
 Example 5; Fig 1; 270pp; English.
 Integrins are a class of membrane-associated molecules that participate
 in cellular adhesion. Integrins are made up of an alpha subunit and a
 beta subunit. One class of human integrins are restricted to expression
 in white blood cells and have a common beta2 subunit: the leukocyte
 integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
 have an important role in immune and inflammatory responses. The present
 protein sequence is the human integrin alpha subunit CD11c. This sequence
 was used in an alignment to identify a novel beta2 integrin alpha
 subunit: alpha d (AA60014 and AA807359). The present sequence has
 approximately 66% identity to the protein sequence of alpha d. The
 Alpha d gene and protein may be useful in therapy for diseases linked to
 aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple
 sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
 distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
 (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the
 inhibition of macrophage infiltration at the site of a central nervous
 system injury. The monoclonal antibodies can also be used to detect and
 diagnose Crohn's disease
 Sequence 1163 AA;
 Query Match 58.6%; Score 3446; DB 3; Length 1163;
 Best Local Similarity 60.8%; Pred. No. 8e-279;
 Matches 686; Conservativity 147; Mismatches 290; Indels 6; Gaps 4;
 1 FNLDENAMTFQENARGFQSVVQLQGRVVVGAQPEIVAANQKSLYQCDYSTGSCBPI 60
 20 FNLDIEELTAFRVDSAGFSDVVQVANSVVVGAQPKIAANQIGLYQCGYSTGACBPI 79
 61 RLQVPVAVNMSLGLSLAATSPOLLACGPTVHQCSTENTYKGLCFPLFQSNLRQOPQK 120
 80 GLQVPPEAVNMSLGLSLASTTSPSOLLACGPTVHQCSTENTYKGLCFPLFQSNLRQOPQK 137
 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDREAKEFISTVMBQLKSKTLFSLMOYSEF 180
 138 LPVSRQECRQCDIVFLIDGSGSISSRNFAFMVFAVRAVISQFQPTQSLAQFSNKF 197
 181 RIHFTKFEQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNTNGARKNAFKILILI 240
 198 QTHFTFEEPRSTNPLSLASVHLQGLQFTYTATQNVVHRLFHASYGARRDAIKILIVI 257
 241 TQKEFGDPLGYEDVTPEDRREGVIRYVIGCDAPFSEKSKQELANTVASKPRDHVQJIN 300
 258 TDGKXGSDLYKVIPEADADAIRIYATGVLAFQNRNMSWKELNDIAKSPQEHIFKVE 317
 301 NFEALNTIQNLREKIFALEGTQTSSSSFEHMSQEGFSAAITSNGFLLSTVGSYDNAG 360
 318 DFDALDKIQNLKEKIFALEGTETISSSSFELEMAQEGFSAVFTPDGPVIGAVGFTWSG 377
 361 GVFLYTSKEKSTFNNTRVDSNDNAYLCYAAAIILRNVRQSLVLTGAPRYQHIGLVAMFR 420
 378 GAFLYPPNMSPTFINNSQENVMRDSYLGSTELALMKGVQSLVLTGAPRYQHIGKAVIFI 437
 421 QNTGMWESNANVKGTOIGVAFGASLCSVDVDSNGSTDLVLICAPHYYEQTRGGQVSCPL 480

Db QVSRQWRKAEVIGTQIGSYFGASLCSVDVDTDLVIGAPHYYEQTRGGQVSCPL 497
 QY PRGORARWODAVLYGBOGQOPWGRFGAALTVLGDVNGDKLTDVAIGAFGEEDNRGAVYLF 540
 Db PRGWR-RWCDVAVLYGBOGQOPWGRFGAALTVLGDVNGDKLTDVWVIGAFGEENRGAVYLF 556
 QY HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVMGAGHVLARSQ 600
 Db HGVLPSPISPSHSQRIAGSKLSPRLQYFGQALSGQDLTQGLVDLAVGARGQVLLRTR 616
 QY PVLRYKAIMFENPREVARNVFECNDQVVKKEAGEVRVCLHVQKSTDRREGOIQSVVT 660
 Db PVLWVGVMGFIPABIPRFAPECRQVVSQTLVQSNICLYIDKRSKNLQSLRDLQSVT 676
 QY YDLALDSGRPHSRVAFNETNSTRRTQVILGLTQTCETLKLQLPNCIEDPVSPVILRLNF 720
 Db LDIALAPGRSLPRAI FQETKNRSLRVRVLGKAHCENFNLLPSCVEDSVIPIILRLNF 736
 QY SLVCTPLSAFENLRPVLAEDAQRALTALFPKKNCGNDNICODDLSTESFMSLDCLVVG 780
 Db TLVGKPLLAFLNRPMLAALAAQRYFTASLPFKKNCADHICQDNMLGISFSPGKLSLLVG 796
 QY GPREFNVTVTRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
 Db SNLENAEVMVWMDGDSYGTITTFSHPAGLSYRYVAEGQKQGLRSLHLC--CSAPVG 854
 QY SGALKSTSCINHPFPENSEVFNITFDVDSKASLGNKLLKANTVSENMMPTNKTEP 900
 Db SQGTWSTSCINHLIFRGAQITFLATFDVSPKAVGLDRLLLIANVSENENIPRTSKTIF 914
 QY QLELPVAVYVYVTVSHGVSTKILNPTAS-ENTSRVMOHOVQVSNLQORSIPLSVLVLP 959
 Db QLELPVAVYVYVTVSHGVSTKILNPTAS-ENTSRVMOHOVQVSNLQORSIPLSVLVLP 974
 QY VRLNQTVIMDRPOVTFSENLSSTCHTKERLPSPSHDFLAELRKAPVNVCSIAVCORIOCDI 1019
 Db VELNQBAVMVMDVEVSHPNPNSLRCSSEKIAPPASDPLAHQKPNVLDSCSIAGCLRFCDV 1034
 QY PFGIOEENFATLKNLSFDWYIKTSHNHLILVSTABILFNDVSFTLLPGQAFVRSQTE 1079
 Db PSFSVQBELDFTLKNLSFGWVROILOKCVSVSVVAELIIFDTSVYSQLPQGEAFRAQTI 1094
 QY TKVFFFPVNPPLPIVGVSSYGGLLALLIATAALYKLGFFKRYQKDMSE 1128
 Db TVLEKYKVHNPPIVGVSSIGLLALLIATAALYKLGFFKRYQKDMSE 1143
 RESULT 11
 ABG61470
 ID ABG61470 standard; protein; 1163 AA.
 XX
 AC ABG61470;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Human Beta2 integrin alphaCD11c subunit.
 XX
 KW Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit; LAD;
 KW leukocyte adhesion deficiency; inflammatory response; diabetes;
 KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW immune complex alveolitis; leukaemia; ICM-R; VCM-1; anti-inflammatory;
 KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
 KW locomotor recovery; locomotor damage; locomotor impairment;
 KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
 OS Homo sapiens.
 XX
 XX WO200230980-A2.
 FN
 XX
 PD 18-APR-2002.
 XX

PF 15-OCT-2001; 2001WO-US032059.
XX
PR 13-OCT-2000; 2000US-00688307.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gallatin WM, Van Der Vieren M;
XX WPI; 2002-463260/49.
XX
XX Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
PT recovery, inhibiting locomotor damage, limiting locomotor impairment, or
PT limiting autonomic and sensory dysfunction following spinal cord injury.
XX
XX Example 5; Page 194-198; 270pp; English.
XX
XX The invention relates to promoting locomotor recovery, inhibiting
CC locomotor damage, limiting locomotor impairment, or limiting autonomic
CC and sensory dysfunction following spinal cord injury by administering an
CC anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a
CC spinal cord injury victim. The method also involves the use of a ligand
CC selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule,
CC vascular cell adhesion molecule). The method is useful for promoting
CC locomotor recovery, inhibiting locomotor damage, limiting locomotor
CC impairment, or limiting autonomic and sensory dysfunction following
CC spinal cord injury. In particular, the spinal cord injury comprises
CC compression of the spinal cord. The antibodies are also useful for
CC reducing inflammation at the site of a central nervous system injury. The
CC specification also details the identification of Beta2 integrin alpha d
CC cDNAs and proteins, for use in raising the antibodies. Beta2 integrins
CC are implicated in diseases such as IAD (leukocyte adhesion deficiency,
CC inflammatory response, diabetes, multiple sclerosis, arthritis, graft
CC atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative
CC colitis, immune complex alveolitis and leukaemia. The present sequence is
CC a Beta2 integrin alpha subunit sequence included for comparison with the
CC Beta2 integrin alpha d protein sequences
XX
XX Sequence 1163 AA;
SQ

Query Match 58.6%; Score 3446; DB 5; Length 1163;
Best Local Similarity 60.8%; Pred. No. 8e-279;
Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;
QY 1 FNLDTEAMTFQENARFGQSVVLQGSRRVVGAPQEIIVAAHQSGSLYQCDYSTGSCPT 60
DB 20 FNLDTEELTAFRVDSAGFGSVVQYANSWVVGAPQKILAAHQIGGLYQGYSTGACEPI 79
QY 61 RLQVFEAVNMSGLSLAATTSPPQLLACGPTVHQTCSNTYVKGCLFLGSLNLQPOPK 120
DB 80 GLQVPPAVNMSGLSLASTTSFQLLACGPTVHHEGGRNMYLTGLCFLLGPT--QLTQR 137
QY 121 PPMALRCQPEDSDIAFLVCGSGSIIIPHDPRRAKEPISTVNEQKKSKTLFSLMOYSEEP 180
DB 138 LPVSRQRCPRQEQDIVFLIDGSGSISRRNFATWNNFVRAVISQFRPSTQFSLMQFSNKF 197
QY 181 RIHETPEFQNNPRLIKPIITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
DB 198 QTHETFEPTRTNPLSLASVHQLQPTTATATQNVVRLPHASVGARDAIKILIVI 257
QY 241 TDGEKFGDPLGYEDVIEADREGVIRYVIGVDAPRSEKSEQLNTVASKPPRDHVPQIN 300
DB 258 TDGKKEGDSLDYKDVIPMAAAGIIRYAIGVGLAFQNRNSWKEINDIASRPSQEHIFKVE 317
QY 301 NFEALKTIQOLREKIFAIETGTGSGSSSEHEMSOEGFSAATSNCPLLSTVSGVDWAG 360
DB 318 DFDALDKIQOLKEKIFAIETGTETISSSSELEMAQEGFSAVFTPDGPVLGAVGVSFTWSG 377
QY 361 GVFLYTSKESKSTFTNMRVDSMDNDAYLGAAYAILNRYQSLVLGAPRYOHIGLVAMER 420
DB 378 GAFILYPPNVSFTFINMSQENVMDSDYLSSTELALWKGVSILVGLAPRYOHIGKAVIFI 437
QY 421 QNTGWESNANVKGTQICATFGASLCSVDVDSNGSITDLVIGAPHYYEQTGRGQSVVCP 480

438 QVSRQWRKAEVIGTQIGSYFGASLCSVDVDTGSDTLVLIGAPHYYEQTGRGQSVVCP 497
481 PRGQARWQCDVLYGEGQPMGRFGAALTVLGDNVNGDKLTDVAIGAPGEENRGAVILF 540
498 PRGMR-RWQDAVLYGEGQPMGRFGAALTVLGDNVNGDKLTDVAIGAPGEENRGAVILF 556
541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDLTVCAGQGHVLLLRQ 600
557 HGVLPSPISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDLTVCAGQGHVLLLR 616
601 PVLRYKAIEMFNPREVARNVFECDQVYKGEAGEVRVCLHVQKTRDLRLRGOIQSVVT 660
617 PVLWVGVSQMFIPASIPRSAFEPCRBQVSEQTLVQSNICLYIDKSKNLLGSRDLQSSVT 676
661 YDLALDSGRPHSRVAVNETKSTRQTOVLGTCETLKLQLPNCIEDPVPVILRLNF 720
677 LDALAPKLSPALFOETKNLSKRVVLGKAHCENFNLLPSCVEDSVPIILRLNF 736
721 SLVGTPLSAPGNLRFVLAEDQRLFTALFPPEKNGCNDNICODDLSITPFSMLDCLVVG 780
737 TLVGKPLLAFLRNLRLPMLAALAQRVFTASLPPEKNGCADHICQDNLGIGSPFGLKSLV 796
781 GPRFNVTYVNDGEDSVRTQVTFPFLDLSYKRVSTLQNRORSORSWELACESASTEV 840
797 SNLELNAEYVWVNDGEDSVGTITTFPSHPAGUSRYVABGQKQGLRSLHLC--CSAPVG 854
841 SGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTTP 900
855 SQGTWSTSCRINHLIFRGAQITFLATEDVSPKAVGLDRLILLIANVSSENNIPRTSKITF 914
901 QLELPKVAIVVTVSHGVSTKYLNFAS-ENTSRVMOHQVQVSNLQBSIPISLVFLVP 959
915 QLELPKVAIVTVSHGVSTKYLNFAS-ENTSRVMOHQVQVSNLQBSIPISLVFLVP 974
960 VRLNQTIVDRPQVTFSENLSSTCHTKERLPQSHSDFLAELRKAAPVNVCSIAVCQRTQCDI 1019
975 VELNQEAVMDVEVSHFPQNPRLCSSEKIAPPASDFLAHQKPNVLDCSIACLFRCDV 1034
1020 PFGQOEPEPNATLKGNLSDWYIKTSHNHLILVSTAEILFNDSTVTLPLPQCAFVSQTE 1079
1035 PSFSVQOEEDFTLKGNLSPGWVRQILQKKVSVVSAEIIFTDSTVZSOLPQCAFVRAQTI 1094
1080 TKVEPEPEPNPLVGVSGVGLLLALITAAALYKLGFFKRYKQMMSE 1128
1095 TVIEKIKVHNPIPLVGVSGVGLLLALITAAALYKLGFFKRYKQMMSE 1143

RESULT 12
ABU07406
ID ABU07406 standard; protein; 1163 AA.
XX AC ABU07406;
XX DT 28-JAN-2003 (first entry)
XX DE Protein differentially regulated in prostate cancer #9.
XX KW Prostate cancer; gene expression; differential regulation;
XX KW molecular marker; drug target; cancer detection; cancer diagnosis;
XX KW cancer staging; cancer grading; cancer assessing; cancer monitoring.
XX OS Homo sapiens.
XX PN WC200281638-A2.
XX XX 17-OCT-2002.
XX XX 08-APR-2002; 2002WO-US010824.
XX XX 06-APR-2001; 2001US-0281731P.
XX XX 06-APR-2001; 2001US-0281732P.
XX XX (ORIG-) ORIGENE TECHNOLOGIES INC.

XX Sun Z, Jay G;
XX WPI; 2003-058520/05.
XX Novel genes which are differentially regulated in prostate cancer, useful
XX for diagnosing prostate cancer in prostate tissue sample and assessing
XX therapeutic or preventive intervention in prostate cancer patients.
XX Claim 1; Page 225-228; 416pp; English.
XX The invention describes genes (I) which are differentially regulated in
XX prostate cancer. (I) is useful for diagnosing a prostate cancer in a
XX sample comprising prostate tissue, which involves determining the number
XX of target genes which are differentially regulated in the sample, where
XX the number is indicative of the probability that the sample comprises
XX prostate cancer. (I) is useful for assessing a therapeutic or preventive
XX intervention in a subject having a prostate cancer, which involves
XX determining the expression levels in a sample comprising prostate tissue
XX of target genes which are differentially regulated in prostate cancer.
XX Preferably, the expression levels of at least 10 genes are determined.
XX (I) is also useful for identifying agents that modulate a biological
XX activity of a polypeptide differentially regulated in prostate cancer
XX cells, which involves contacting a polypeptide differentially-regulated
XX in prostate cancer cells with a test agent under conditions effective for
XX the test agent to modulate a biological activity of the polypeptide, and
XX determining whether the test agent modulates the biological activity. (I)
XX is useful as molecular markers, as drug targets, and for detecting,
XX diagnosing, staging, grading, assessing, monitoring, prognosticating,
XX preventing or treating, determining predisposition to diseases and
XX conditions especially relating to prostate cancer. (I) and its expression
XX products are used in the diagnostic test to assay for presence of cancer
XX e.g., in tissue sections, in biopsy samples, in total RNA, in lymph, in
XX blood etc. (I) is useful for assessing cancer e.g., to determine the type
XX of cancer, its stage of development, the nature of genetic defect, etc.
XX The polypeptide encoded by (I) can be used as target for therapy or drug
XX discovery. (I) can also be used for expressing the polypeptide and thus
XX for searching specific binding partners of the polypeptide. (I) is useful
XX in therapeutic applications to treat prostate cancer. The identification
XX of specific genes, and groups of genes, expressed in pathways
XX physiologically relevant to prostate cancer permits the definition of
XX functional and disease pathways and the delineation of targets in these
XX pathways which are useful in diagnostic, therapeutic, and clinical
XX applications. This is the amino acid sequence of a protein differentially
XX regulated in prostate cancer
XX
XX Sequence 1163 AA;
Query Match 58.6%; Score 3444; DB 6; Length 1163;
Best Local Similarity 60.7%; Pred. No. 1.2e-278;
Matches 585; Conservative 139; Mismatches 299; Indels 6; Gaps 4;
QY 1 FNLDTEMTAFQENARGFGQSVVLOGSRVWVGAPOEIVAAORGSLYOCXYSTGSCPT 60
DB 20 FNLDTEMTAFQENARGFGQSVVLOGSRVWVGAPOEIVAAORGSLYOCXYSTGSCPT 79
QY 61 RLQVPEAVNMISGLSLAATTSPPQLACGPTVHQTCSNTYVKGCLFLGSLNLRQPOK 120
DB 80 GLQVPEAVNMISGLSLAATTSPPQLACGPTVHQTCSNTYVKGCLFLGSLNLRQPOK 137
QY 121 FPEALRGCPQEDSDIAFLVDGSGISLTHDFPRAKEFISTVMEQLKSKTLFSLMOYSEEP 180
DB 138 LPVSRQECPRQEDSDIAFLVDGSGISLTHDFPRAKEFISTVMEQLKSKTLFSLMOYSEEP 197
QY 181 REHFTPEFQNNPNSRLKIPITQLLCRTHATGIRKVRRELFTNITNGARKNAFKILILI 240
DB 198 QTHLTFEFTRTNPNLSLASVHLQGFYTTATAIQNVVHLLFHASVGAARDATKILIVI 257
QY 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAFSEKSEKQELNTVASKPRDHFQIN 300
DB 258 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAFSEKSEKQELNTVASKPRDHFQIN 317
QY 301 NFEALTKIQNLREKIPITQLLCRTHATGIRKVRRELFTNITNGARKNAFKILILI 360

DB 318 DFDALNDIQTLREKIPITQLLCRTHATGIRKVRRELFTNITNGARKNAFKILILI 377
QY 361 GVELYTSKEKSTFINNTRVDSMDNDAYLGYAAAILIIRNRVQSILVGLGAPRYOHLVAMPR 420
DB 378 GAFLYPNNSPTFINNTRVDSMDNDAYLGYAAAILIIRNRVQSILVGLGAPRYOHLVAMPR 437
QY 421 QNTGMESNANVKTGTOIGAVFGSLCSVDVDSNGSTDLVLIGAPHYHYBOYRGQVSCPL 480
DB 438 QVSRQWRMKAETVGTQIGSYFGSLCSVDVDSNGSTDLVLIGAPHYHYBOYRGQVSCPL 497
QY 481 PRGORAFWQCDAYLGEQSGQWGRFGAALTVLGDVNGDKLTVAIGAPGEDNRGAYILF 540
DB 498 PRGWR-RWKCDAVLYGEQSGQWGRFGAALTVLGDVNGDKLTVAIGAPGEDNRGAYILF 556
QY 541 HGTSGSISSESHQRIAGSLSPRLQYFGOSLGGDLTWDGLVLDLTVGAQGHVLLRSO 600
DB 557 HGVLPSPISSESHQRIAGSLSPRLQYFGOSLGGDLTWDGLVLDLTVGAQGHVLLRSO 616
QY 601 PVLKVKXIMEFNRVARNVFECDQVVKGEAGEVRCVCLHVQKSTRDRLEBOIQSVVT 660
DB 617 PVLWVGYSMQFIPABIPRSAFECEQVWSEQLTVQSNICLVIDKRSKNLLGSRDLQSSVT 676
QY 661 YDLALDSGRPHSRVAVNTRKSTRQVGLTGTCTETKLOLPNCIEDPVSPTVLRNF 720
DB 677 LDALDPGLSPRATFOETKRSLSRVRLGLKAHCENFNLLPSCVEDSVPTITLRNF 736
QY 721 SLVGTPLSAFNTLRPLVLAEDAQRFTALFPEKNCNDNICQDDLSTFSSMSLCLVVG 780
DB 737 TLVQKPLILAFNLRPLVLAEDAQRFTALFPEKNCNDNICQDDLSTFSSMSLCLVVG 796
QY 781 GPREFNVTVVRNDGDSYRTQVTFPPDLSTLYKRVKSTLQNRQSRWRLACBSASTEV 840
DB 797 SNLEINAEVWVNDGDSYRTQVTFPPDLSTLYKRVKSTLQNRQSRWRLACBSASTEV 854
QY 841 SGALKSTSCSNHDPFENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 900
DB 855 SQGTWSTSCRINHLIFRGAGAITLATFDSPKAVLGRLLLTANVSENTPRTSKTF 914
QY 901 QLELPVKYAVTWVYTSQVSTKYLNTAS-ENTSRVMOHQYQVSNLGRSIPISLVFLVP 959
DB 915 QLELPVKYAVTWVYTSQVSTKYLNTAS-ENTSRVMOHQYQVSNLGRSIPISLVFLVP 974
QY 960 VRLNQTVINDRPQVTFSENLSSTCHTKEELSHSDPLAELKAPVWNCISAVCORICDI 1019
DB 975 VELNQAVVMDVEVSLPQNSLCSSEKIAQSPADFLAHIQNPVLDGSIAGCLFRCDV 1034
QY 1020 PFGIOEEFNATLKNLSFDWVWIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFRSOTE 1079
DB 1035 PPSVQOEELDTLKNLSFGVVRQLQKKVSVVAEITFDTSVYSQLPQGEAFRAQTT 1094
QY 1080 TKVEPFVNPPLIVGSSVGLLILALITAAIKLGPKEQYKDMMS 1128
DB 1095 TVLEKVKVHNPTPLIVGSSIGGLLILALITAVLYKVGPFKQYKEMMBE 1143
RESULT 13
AA78166
ID AA78166 standard; protein; 1161 AA.
XX
AC AA78166;
XX
DT 28-DEC-1995 (first entry)
XX
DE Human beta-2 integrin alpha-d.
XX
KW Beta-2 integrin alpha-d subunit; antinflammatory; arteriosclerosis;
XX inflammatory bowel disease; asthma.
XX
OS Homo sapiens.
XX
FH Key
FT Domain 17..1108

FT FT Region /note= "extracellular domain"
 FT 150..352
 FT /note= "this region is homologous to the insert common to
 FT Ch1a,b,c and may be a site for interaction with ICM
 FT family proteins"
 FT Binding-site 465..474
 FT Binding-site 518..527
 FT /note= "putative cation binding site"
 FT Binding-site 592..600
 FT /note= "putative cation binding site"
 FT Binding-site 1109..1128
 FT /note= "transmembrane region"
 FT Domain 1129..1161
 FT /note= "cytoplasmic domain"
 PN W09517412-A1.
 XX
 XX 29-JUN-1995.
 XX 21-DEC-1994; 94WO-US014832.
 XX 23-DEC-1993; 93US-00173497.
 PR 05-AUG-1994; 94US-00286889.
 XX (ICOS-) ICOS CORP.
 PA Gallatin WM, Van Der Vieren M;
 PI WPI; 1995-240603/31.
 XX N-PSDB; AAQ91712.
 DR Alpha sub-unit polypeptide of human beta 2 integrin - used to identify
 PT potential antiinflammatory agents, for the treatment of graft
 PT arteriosclerosis, inflammatory bowel disease, asthma, etc.
 FT
 XX Claim 7; Page 82-87; 172pp; English.
 CC A probe based on a partial cDNA clone (given in AAQ91727) of canine alpha
 CC -TM1 was used to screen a human spleen cDNA library to identify clone
 CC 19A2 encoding the beta-1 integrin alpha-subunit, alpha-d. The cDNA was
 CC manipulated to allow expression of recombinant alpha-d subunit in COS and
 CC CHO Cells
 XX
 XX Sequence 1161 AA;
 SQ
 Query Match 58.0%; Score 3411; DB 2; Length 1161;
 Best Local Similarity 59.4%; Pred. No. 6.8e-276; Indels 8; Gaps 6;
 Matches 671; Conservative 166; Mismatches 284;
 QY 1 FNLDTENAMTFQENARFGQSVVQLQGSRVVVGAPQEIIVAAQNGSLYQCYSTGSCPEI 60
 DB 17 FNLDVEPTIFQEDAGFGQSVVQFGSRLVVGAPLEVVAANQTRGLYDCAAAATGWCQPI 76
 QY 61 RLOVFEAVNMSLGLSLAATTSPQLLACGPTVHTQCSNTYVKGCLPLFGSNLQPOPK 120
 DB 77 PLHIREAVNMSLGLSLAATTSPQLLACGPTVHTQCSNTYVKGCLPLFGSNLQPOPK 135
 QY 121 FPEALRGCPQEDDIAFLVDGSGIIPHPDRAKEPISTVMEOLKSKTLFSLMAYSEEF 180
 DB 136 VPDATPECHQENDIVFLDGGSSIDONDENQMGFQVQVMOQFEGDTLFLMAYSNLL 195
 QY 181 RHFTPEFQNNPNRSLKIPITQLLGRTHATGIRKVRRELPNITNGARKNAFKILLI 240
 DB 196 KHFTTQFTSPSQSLVDPIVLQGLTFTATGILFVVTQLFHHKNGARKSAKKILVI 255
 QY 241 TDEKEFGDPLGYEDVPEADREGVIRYVIGWDAPFSEKSRQBLNTVASKPPRDRHVQIN 300
 DB 256 TDQCKYKDPLEYSDVLPQEKAGIIRYGVGHAFQGPATARQELNITSSAPQDHVFKYD 315
 QY 301 NFPEALTYIQNLREKIFAIGTGTGSSSFHEHMSQEGFSAITSNGLPILLTVGSDWAG 360
 DB 316 NFAALGSIQKQLEKIVAVEGTQSRASSSPQHEMSQEGFSTALTMGDLGLGAVGFSWSG 375

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
 DB 376 GAFLYPPNMSPTFINMSQENVDMDSYLGYSTELALWKGQVNLVGLGAPRYQHTKAVFT 435
 QY 421 QNTGWESNANVKTQTOIGAYFGASLCSVDVDSNGSTDLVILGAPHYHQTGGQSVVCPL 480
 DB 436 QVSRQMKCAEVGTQIGSYFGASLCSVDVDSNGSTDLVILGAPHYHQTGGQSVVCPL 495
 QY 481 PRQARWQCDVAVLYGQGPWGRFGAALTIVLGVNDGDKLTVDVAIGAPGEDNAGVYLF 540
 DB 496 PRQVQWQCDVAVLYGQGPWGRFGAALTIVLGVNDGDKLTVDVAIGAPGEDNAGVYLF 555
 QY 541 HGTSGGISPSHSORIASKLSPLRQYFGQSLGQDLTMDCLVDLTVAGAGHVLRSO 600
 DB 556 HGASEGISPSHSORIASKLSPLRQYFGQSLGQDLTMDCLVDLTVAGAGHVLRSO 615
 QY 601 PVLRVKALMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 660
 DB 616 PVLKVGVMRFPVEVAKAVYRWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673
 QY 661 YDLALDSGRPHSRVAVFNETKNSRTOVGLTCTETLKLQPNCTEDPVSPIVLRNF 720
 DB 674 FDLALDPGLRTSRAIFNETKNPTLTRKTLGLGHCTETLKLQPNCTEDPVSPIVLRNF 733
 QY 721 SLVGTPLSAPGNLRPLAEDAQREPTALFPEKNCGMNDIQQDLSITFSFMSLDCLVVG 780
 DB 734 SLVREPSPQNLAPVLAGSQDLPTASLPEKNCQDGLCEGDLGVTLSFGLQTLTVG 793
 QY 781 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYKVKSTLQNRQSRWRLACESASSTEV 840
 DB 794 SSLELVTVVW 852
 QY 841 SGALKSTSCSNHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENMPDNKTEF 900
 DB 853 EG-LRSSKCSVNHPIFHEGSGTIFIVFDVSKATLGDRLMARASSENKASSSKATP 911
 QY 901 QLELPVKYAVTVVTSVHGVSTKYNLF-TASENTRVMOHQYOVNLSLQORSIPISLVFLVP 959
 DB 912 QLELPVKYAVTVVTSVHGVSTKYNLF-TASENTRVMOHQYOVNLSLQORSIPISLVFLVP 971
 QY 960 VRLNQTVWDRPQVTPPSNLSTCTKTERLPSHSDPLAELKAPVWNCISIAVCORIQCDI 1019
 DB 972 VLLNGVAVW 1029
 QY 1020 PFFGIQEEFNATLKGNSLFDWYIKTSENHLLIVTAEILFNDVSFTLLPGQAFVRSQTE 1079
 DB 1030 PFSVQSELDFTLKGNSLFGVWRETLOKVLVSVVAEITPDTSVVSQLPGQAFVRSQTE 1089
 QY 1080 TKVPEFVFNPLDIVSGVSGILLILITAAALYKLGFFKRYKMKSE 1128
 DB 1090 MVLEDEYVNAIPTIMGSSVGAALLLITATLYKLGFFKRYKMKSE 1138
 RESULT 14
 AAW23049
 ID AAW23049 standard; protein; 1161 AA.
 XX
 AC AAW23049;
 XX
 DT 24-FEB-1998 (first entry)
 XX
 DE Human beta 2 integrin alpha d subunit.
 XX
 XX Beta 2 integrin alpha d subunit; human; cell migration; cell adhesion;
 KW phagocytosis; diabetes; atherosclerosis; multiple sclerosis; asthma;
 KW psoriasis; lung inflammation; acute respiratory distress syndrome;
 XX rheumatoid arthritis.
 OS Homo sapiens.
 XX Key location/Qualifiers
 FH 17..1108
 FT Domain

FT	Domain	/label= Extracellular_domain 150..352 /note= "region homologous to the I (insertion) domain common to CD11a, CD11b and CD11c" 1109..1128 /label= Transmembrane_domain /note= "homologous to the human CD11c transmembrane region" 1129..1161 /label= Cytoplasmic_domain
FT	XX	WO9731099-A1.
PN	XX	28-AUG-1997.
PD	XX	24-FEB-1997; 97WO-US002713.
PP	XX	22-FEB-1996; 96US-00605672.
XX	XX	(ICOS-) ICOS CORP.
XX	XX	Gallatin WM, Van Der Vieren M;
PI	XX	WPI; 1997-435154/40.
DR	XX	N-PSDB; AAT79220.
XX	XX	Hybridoma 199M and antibody secreted by it - specific for new rat beta2 integrin subunit, useful to detect subunit in cells and modulate its activity.
PT	XX	Example 5; Page 116-120; 22pp; English.
PT	XX	This polypeptide comprises a novel human beta 2 integrin subunit, designated alpha d. Its sequence was deduced from a cDNA clone (see AAT79220) isolated from a spleen cDNA library. Alpha d is involved in cell migration, phagocytosis and cell-cell interaction. Recombinant alpha d polypeptides can be expressed in transformed host cells for use in assays for identifying antibodies or other compounds that modulate alpha d activity or which modulate the interaction between alpha d and a ligand, for treating or preventing diseases in which macrophages are implicated. Treatment is applicable to disease states in which alpha d binding, or localised accumulation of cells which express alpha d, is implicated such as such as type I diabetes, arteriosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome and rheumatoid arthritis
CC	XX	Sequence 1161 AA;
SQ	XX	Query Match 58.0%; Score 3411; DB 2; Length 1161; Best Local Similarity 59.4%; Pred No 6,8e-276; Matches 671; Conservative 166; Mismatches 284; Indels 8; Gaps 6;
QY	1	FNLDTENAMTFQENARGFGQSVMVQLQGSRVVVGAPQEIVANQRGSLYQCDDYSTGSCEPI 60 17 FNLDEEPTIFQEDAGGFQSVVQGGSRVLVVGAPLEVVAANTGRLYDCAAAATGMCPPI 76
Db	61	RLOVPVEANNMSLGLSLAATSPPOLACGPTHVTCSENTYVVKGLCFGLGNLRQQPOK 120 77 PLHIREPAVNMSLGTLAASINGSLACGPTLHRVCGENSYSGSCLLIGSRW-EIIQT 135
QY	121	PFEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEPTSTVMEQLKKSKTILFSLMOYSSEEF 180 136 VPDAITECHPHOEMDIVFLIDGSSIDQNDEMCKMGFEVQAQVQEGTDITLFALMOYSNEL 195
QY	181	RHPHFKEFNQNNPRSLIKPTQLGLGTHHTATGTRKVVRLEFNITGARKNAFKILLI 240 196 KIHFTPTQRTSPSQSLVDPIVLQKGITFTATGTLTVTLQFHNGARKSAKKILLIVI 255
QY	241	TDCKEKGDLDGYEDIPEADREGVIRYVGVGDAPRSEKSEQELNLTAVSKPPDPDHVFQIN 300 256 TDGQKYKDPLEYSDVTPQAKAGIRYAIVGHAFQGTARQELNTISSAPPQDHVFKVD 315
QY	301	NFEALKTIONQREKIFALETGQTGSSSFHEHMQBGSAAITNSGPLISTVGSYDWAG 360

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:16:51 ; Search time 38.8905 Seconds
(without alignments)
8225.189 Million cell updates/sec

Title: US-09-902-481B-5
Perfect score: 5876
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKQYKDMSEGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 28138677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgm2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgm2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgm2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgm2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgm2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgm2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgm2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgm2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgm2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgm2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgm2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgm2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgm2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgm2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgm2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgm2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgm2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgm2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5876	100.0	1137	10 US-09-902-481A-5	Sequence 5, Appli
2	5862	99.8	1153	9 US-09-350-259-3	Sequence 3, Appli
3	5862	99.8	1153	10 US-09-902-481A-1	Sequence 1, Appli
4	5862	99.8	1153	10 US-09-891-943-3	Sequence 3, Appli
5	5862	99.8	1153	14 US-10-144-259-30	Sequence 176, App
6	5862	99.8	1153	14 US-10-207-655-176	Sequence 6, Appli
7	5855	99.6	1137	10 US-09-902-481A-6	Sequence 4, Appli
8	5851	99.6	1137	10 US-09-902-481A-4	Sequence 4, Appli
9	5846.5	99.5	1152	9 US-09-945-265-4	Sequence 3, Appli
10	5836	99.3	1137	10 US-09-902-481A-3	Sequence 204, App
11	3469	59.0	1163	15 US-10-116-275-204	Sequence 4, Appli
12	3446	58.6	1163	9 US-09-350-259-4	Sequence 4, Appli
13	3446	58.6	1163	10 US-09-891-943-4	Sequence 2, Appli
14	3411	58.0	1161	9 US-09-350-259-2	Sequence 2, Appli
15	3411	58.0	1161	10 US-09-891-943-2	Sequence 2, Appli

16	3395.5	57.8	1161	9 US-09-350-259-99	Sequence 99, Appli
17	3395.5	57.8	1161	10 US-09-891-943-99	Sequence 99, Appli
18	3332.5	55.0	1161	9 US-09-350-259-55	Sequence 55, Appli
19	3332.5	55.0	1161	10 US-09-891-943-55	Sequence 55, Appli
20	3222	54.8	1161	9 US-09-350-259-53	Sequence 53, Appli
21	3222	54.8	1161	10 US-09-891-943-53	Sequence 53, Appli
22	3211.5	54.7	1151	9 US-09-350-259-37	Sequence 37, Appli
23	3211.5	54.7	1151	10 US-09-891-943-37	Sequence 37, Appli
24	3201	54.5	1155	9 US-09-350-259-46	Sequence 46, Appli
25	3201	54.5	1155	10 US-09-891-943-46	Sequence 46, Appli
26	1848	31.4	369	12 US-10-087-192-1212	Sequence 1212, Ap
27	1557.5	26.5	1170	9 US-09-945-265-2	Sequence 2, Appli
28	1553.5	26.4	1170	12 US-10-261-164-1	Sequence 1, Appli
29	1530	26.0	1223	16 US-10-408-765A-295	Sequence 295, App
30	1359.5	23.1	1086	16 US-10-408-765A-1871	Sequence 1871, Ap
31	1229.5	20.9	494	9 US-09-350-259-103	Sequence 103, App
32	1229.5	20.9	494	10 US-09-891-943-103	Sequence 103, App
33	1153	19.6	1179	14 US-10-177-550-2	Sequence 2, Appli
34	1153	19.6	1179	15 US-10-173-551-2	Sequence 101, App
35	1151.5	19.6	413	9 US-09-350-259-101	Sequence 101, App
36	1151.5	19.6	413	10 US-09-891-943-101	Sequence 101, App
37	1103.5	18.8	1151	10 US-09-984-130-103	Sequence 103, App
38	1103.5	18.8	1151	10 US-09-836-353A-103	Sequence 103, App
39	1103.5	18.8	1179	12 US-09-918-715-250	Sequence 250, App
40	1101	18.7	1188	15 US-10-291-265-810	Sequence 810, App
41	1097	18.7	1188	15 US-10-291-265-338	Sequence 338, App
42	1093.5	18.6	1189	10 US-09-984-130-35	Sequence 35, Appli
43	1093.5	18.6	1189	10 US-09-836-353A-35	Sequence 35, Appli
44	1093.5	18.6	1189	12 US-10-262-839-4	Sequence 4, Appli
45	1084	18.4	589	12 US-10-261-164-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-902-481A-5
; Sequence 5, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-5

Query Match	100.0%;	Score 5876;	DB 10;	Length 1137;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1137;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	FNLDTENAMTFQENARGFGQSVVOLQSGRVVVGAPQIIVANQSGLSYQCDYSTGSCPEI	60	
Db	1	FNLDTENAMTFQENARGFGQSVVOLQSGRVVVGAPQIIVANQSGLSYQCDYSTGSCPEI	60	
Qy	61	RLOVPVAVNMSLGLSLAATTSPQLACGPTVHTCTSENTYVKGCLCFPGSNLRQPOK	120	
Db	61	RLOVPVAVNMSLGLSLAATTSPQLACGPTVHTCTSENTYVKGCLCFPGSNLRQPOK	120	
Qy	121	FFPALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFTSTVMEQLKXKTLPSLMQSEEP	180	

Db	121	FPFALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFTSTVMEQKKSKTLFSLMQSSEF	180
Qy	181	RIHFTPEFONNPNRSLIKPIITQLGRTHATGIRKVVRELNTNGARKNAFKILILI	240
Db	181	RIHFTPEFONNPNRSLIKPIITQLGRTHATGIRKVVRELNTNGARKNAFKILILI	240
Qy	241	TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQELNTVASKPRDHVQIN	300
Db	241	TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQELNTVASKPRDHVQIN	300
Qy	301	NFEALKTIONOLREKIPALTEGTGSSSPHEHMSQEGFSAATNSGPLLSTVGSYDWAQ	360
Db	301	NFEALKTIONOLREKIPALTEGTGSSSPHEHMSQEGFSAATNSGPLLSTVGSYDWAQ	360
Qy	361	GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIIILNRVQSLVGLAPRYOHIGLVAMR	420
Db	361	GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIIILNRVQSLVGLAPRYOHIGLVAMR	420
Qy	421	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIIGAPHYVETRGQVSCPL	480
Db	421	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIIGAPHYVETRGQVSCPL	480
Qy	481	PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF	540
Db	481	PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF	540
Qy	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDPLTMDGLVDLTVGAQGHVLLRSQ	600
Db	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDPLTMDGLVDLTVGAQGHVLLRSQ	600
Qy	601	PVLRVKAIEMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	660
Db	601	PVLRVKAIEMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	660
Qy	661	YDLALDSGRPHSAVFNENKSTRQTVLGLTQTCETKLQLPNCIEDPVSPIVLRNLF	720
Db	661	YDLALDSGRPHSAVFNENKSTRQTVLGLTQTCETKLQLPNCIEDPVSPIVLRNLF	720
Qy	721	SLVGTPLSAFNLAPVLAEDAQRLLFTALFPPEKNCNDNICODDLSITFSFMSLDCILVWG	780
Db	721	SLVGTPLSAFNLAPVLAEDAQRLLFTALFPPEKNCNDNICODDLSITFSFMSLDCILVWG	780
Qy	781	GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYKVKSTLQNRQSQRWRLACSSASTEV	840
Db	781	GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYKVKSTLQNRQSQRWRLACSSASTEV	840
Qy	841	SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF	900
Db	841	SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF	900
Qy	901	QLELPVKYAVVMVTSYGVSTKYLNFTASENTSRVMQHQYQVSNLQORSPLISLVFLVPV	960
Db	901	QLELPVKYAVVMVTSYGVSTKYLNFTASENTSRVMQHQYQVSNLQORSPLISLVFLVPV	960
Qy	961	RINQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVVNCISIAVCQRIQDIP	1020
Db	961	RINQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVVNCISIAVCQRIQDIP	1020
Qy	1021	PFQIQEENATLKNLSFDWIKTSHNHLIIVSTAEILFNDSTVFTLLPGQAGFVRSQTET	1080
Db	1021	PFQIQEENATLKNLSFDWIKTSHNHLIIVSTAEILFNDSTVFTLLPGQAGFVRSQTET	1080
Qy	1081	KVEPFEPVNPPLIVGSGVGLLILALITAALYKLGFFKQYKDMXSEGGPPGAEPO	1137
Db	1081	KVEPFEPVNPPLIVGSGVGLLILALITAALYKLGFFKQYKDMXSEGGPPGAEPO	1137

RESULT 2
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. US2002062008A1
; GENERAL INFORMATION:

APPLICANT: Gallatin, Michael W.
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. US2002062008A1el Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/350,259
CURRENT FILING DATE: 1999-07-08
EARLIER APPLICATION NUMBER: 09/193,043
EARLIER FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: 08/173,497
EARLIER FILING DATE: 1993-12-23
EARLIER APPLICATION NUMBER: 08/286,889
EARLIER FILING DATE: 1994-08-05
EARLIER APPLICATION NUMBER: 08/362,652
EARLIER FILING DATE: 1994-12-21
EARLIER APPLICATION NUMBER: 08/943,363
EARLIER FILING DATE: 1997-10-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 1153
TYPE: PRT
ORGANISM: Homo sapiens
US-09-350-259-3

Qy	1	FNLDTENAMTFOENARGFGQSVVQLOGSRVVVGAPQEIIVANQORSLYQCDYSTGSCPEI	60
Db	17	FNLDTENAMTFOENARGFGQSVVQLOGSRVVVGAPQEIIVANQORSLYQCDYSTGSCPEI	76
Qy	61	RLOQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVVKGLCFPGSNLRQPOPK	120
Db	77	RLOQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVVKGLCFPGSNLRQPOPK	136
Qy	121	FPFALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFTSTVMEQKKSKTLFSLMQSSEF	180
Db	137	FPFALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFTSTVMEQKKSKTLFSLMQSSEF	196
Qy	181	RIHFTPEFONNPNRSLIKPIITQLGRTHATGIRKVVRELNTNGARKNAFKILILI	240
Db	197	RIHFTPEFONNPNRSLIKPIITQLGRTHATGIRKVVRELNTNGARKNAFKILILI	256
Qy	241	TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQELNTVASKPRDHVQIN	300
Db	257	TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQELNTVASKPRDHVQIN	316
Qy	301	NFEALKTIONOLREKIPALTEGTGSSSPHEHMSQEGFSAATNSGPLLSTVGSYDWAQ	360
Db	317	NFEALKTIONOLREKIPALTEGTGSSSPHEHMSQEGFSAATNSGPLLSTVGSYDWAQ	376
Qy	361	GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIIILNRVQSLVGLAPRYOHIGLVAMR	420
Db	377	GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIIILNRVQSLVGLAPRYOHIGLVAMR	436
Qy	421	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIIGAPHYVETRGQVSCPL	480
Db	437	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIIGAPHYVETRGQVSCPL	496
Qy	481	PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF	540
Db	497	PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF	556
Qy	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDPLTMDGLVDLTVGAQGHVLLRSQ	600
Db	557	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDPLTMDGLVDLTVGAQGHVLLRSQ	616
Qy	601	PVLRVKAIEMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	660
Db	617	PVLRVKAIEMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	676
Qy	661	YDLALDSGRPHSAVFNENKSTRQTVLGLTQTCETKLQLPNCIEDPVSPIVLRNLF	720

```

Db      677 YDLDSDGRPHSRVAFNETKNSRTRQVGLGLTQTCETLKLQPNCTIEDPVSIVLRLNF 736
Qy      721 SLVGTPLSAFGLNRPVLAEDAQRFTALPPEKNCNDNICODDLSITPFSFMSLDCLVWG 780
Db      737 SLVGTPLSAFGLNRPVLAEDAQRFTALPPEKNCNDNICODDLSITPFSFMSLDCLVWG 796
Qy      781 GPRFNVTVTVRNDEGDSYRTQVTFPPDLQSVKSTLQNSQSRSLWRLACSSASTEV 840
Db      797 GPRFNVTVTVRNDEGDSYRTQVTFPPDLQSVKSTLQNSQSRSLWRLACSSASTEV 856
Qy      841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEP 900
Db      857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEP 916
Qy      901 QLELPVKYAVVMVTVSHGVSTKYLNFNTASENTSRVMQHQYQVSNLQORSLSPLSLVFLVPV 960
Db      917 QLELPVKYAVVMVTVSHGVSTKYLNFNTASENTSRVMQHQYQVSNLQORSLSPLSLVFLVPV 976
Qy      961 RLNQTVIWDREPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVNCSIAVCORIOCDIP 1020
Db      977 RLNQTVIWDREPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVNCSIAVCORIOCDIP 1036
Qy      1021 PFGIOEFNATLKGNSLDFDWMYIKTSHNHLIVSTAEILFNDSVFTLLPGQCAFVRSOTET 1080
Db      1037 PFGIOEFNATLKGNSLDFDWMYIKTSHNHLIVSTAEILFNDSVFTLLPGQCAFVRSOTET 1096
Qy      1081 KVPEPEVNPPLIVGSSVGLLALLITAAALYKLGFFKQYKXMMSEGGPPGAEPO 1137
Db      1097 KVPEPEVNPPLIVGSSVGLLALLITAAALYKLGFFKQYKXMMSEGGPPGAEPO 1153

```

RESULT 3

```

US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimada, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (17)..()
; OTHER INFORMATION:

```

```

Query Match          99.8%; Score 5862; DB 10; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy      1  FNLDTENAMTFOENARGFGQSVVQSGSRVVGAPQBIIVAAHQSGLYQCDYSTGSCBPI 60
Db      17  FNLDTENAMTFOENARGFGQSVVQSGSRVVGAPQBIIVAAHQSGLYQCDYSTGSCBPI 76

Qy      61  RLQVPEAVNMSGLSLAATTSPPQLACGPTTHQTCSENTYKGLCFPLGSLRQOPK 120
Db      77  RLQVPEAVNMSGLSLAATTSPPQLACGPTTHQTCSENTYKGLCFPLGSLRQOPK 136

Qy      121 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMEKFVSVTVMQKSKTKLFLSLMOYSEEF 196

```

```

Db      137 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMEKFVSVTVMQKSKTKLFLSLMOYSEEF 196
Qy      181 RIHFTFEPQNNPNSRSLIKPITCOLLGRTHATGIRKVVRELEFNITWGARNAPKILLI 240
Db      197 RIHFTFEPQNNPNSRSLIKPITCOLLGRTHATGIRKVVRELEFNITWGARNAPKILLI 256
Qy      241 TDGKFGDPLGYEYDVIPEADREGVIRYVIGVGDAPFRSEKSRQELMTASKPPRDHVFQIN 300
Db      257 TDGKFGDPLGYEYDVIPEADREGVIRYVIGVGDAPFRSEKSRQELMTASKPPRDHVFQIN 316
Qy      301 NFEALKTIONOLREKIIPAIETOTGSSSSFEHMSORGPSAAITNSGPLSTVGSYDWDAG 360
Db      317 NFEALKTIONOLREKIIPAIETOTGSSSSFEHMSORGPSAAITNSGPLSTVGSYDWDAG 376
Qy      361 GVFLYTSSEKSTFNTMTRVDSMDNDAYLVAAAILLRNVOSVLGAPRYOHICLVAMFR 420
Db      377 GVFLYTSSEKSTFNTMTRVDSMDNDAYLVAAAILLRNVOSVLGAPRYOHICLVAMFR 436
Qy      421 QNTGMBESNANVKGTQIAGYFASLCSVDVDSNGSTDVLIGAPHYYEOTRGQSVSCPL 480
Db      437 QNTGMBESNANVKGTQIAGYFASLCSVDVDSNGSTDVLIGAPHYYEOTRGQSVSCPL 496
Qy      481 PRGORARWQCDVLYGREGQGWGFGAALTIVLGDVNGDKLTDVAIGAPGBEDNRGAVYLF 540
Db      497 PRGORARWQCDVLYGREGQGWGFGAALTIVLGDVNGDKLTDVAIGAPGBEDNRGAVYLF 556
Qy      541 HGTSGSGISPSHSORLAGSKLSPRLQYFGOSLGGQDLTWDGHLVDTLVGAQGHVLLRSQ 600
Db      557 HGTSGSGISPSHSORLAGSKLSPRLQYFGOSLGGQDLTWDGHLVDTLVGAQGHVLLRSQ 616
Qy      601 PVLRYKAIMFNPREVARNVFECDQVVKGKAGEVEVCLHVQKSTRDRLRBQIQSVVT 660
Db      617 PVLRYKAIMFNPREVARNVFECDQVVKGKAGEVEVCLHVQKSTRDRLRBQIQSVVT 676
Qy      661 YDLALDSGRPHSRVAFNETKNSRTRQVGLGLTQTCETLKLQPNCTIEDPVSIVLRLNF 720
Db      677 YDLALDSGRPHSRVAFNETKNSRTRQVGLGLTQTCETLKLQPNCTIEDPVSIVLRLNF 736
Qy      721 SLVGTPLSAFGLNRPVLAEDAQRFTALPPEKNCNDNICODDLSITPFSFMSLDCLVWG 780
Db      737 SLVGTPLSAFGLNRPVLAEDAQRFTALPPEKNCNDNICODDLSITPFSFMSLDCLVWG 796
Qy      781 GPRFNVTVTVRNDEGDSYRTQVTFPPDLQSVKSTLQNSQSRSLWRLACSSASTEV 840
Db      797 GPRFNVTVTVRNDEGDSYRTQVTFPPDLQSVKSTLQNSQSRSLWRLACSSASTEV 856
Qy      841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEP 900
Db      857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEP 916
Qy      901 QLELPVKYAVVMVTVSHGVSTKYLNFNTASENTSRVMQHQYQVSNLQORSLSPLSLVFLVPV 960
Db      917 QLELPVKYAVVMVTVSHGVSTKYLNFNTASENTSRVMQHQYQVSNLQORSLSPLSLVFLVPV 976
Qy      961 RLNQTVIWDREPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVNCSIAVCORIOCDIP 1020
Db      977 RLNQTVIWDREPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVNCSIAVCORIOCDIP 1036
Qy      1021 PFGIOEFNATLKGNSLDFDWMYIKTSHNHLIVSTAEILFNDSVFTLLPGQCAFVRSOTET 1080
Db      1037 PFGIOEFNATLKGNSLDFDWMYIKTSHNHLIVSTAEILFNDSVFTLLPGQCAFVRSOTET 1096
Qy      1081 KVPEPEVNPPLIVGSSVGLLALLITAAALYKLGFFKQYKXMMSEGGPPGAEPO 1137
Db      1097 KVPEPEVNPPLIVGSSVGLLALLITAAALYKLGFFKQYKXMMSEGGPPGAEPO 1153

```

RESULT 4

```

US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US2003007728A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.

```

APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-3

Query Match 99.8%; Score 5862; DB 10; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAAQNGSLYQCDYSTGSCBPI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAAQNGSLYQCDYSTGSCBPI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFLFGSNLRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFLFGSNLRQPOK 136

QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEPISTVMEOLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEPISTVMEOLKSKTFLSLMOYSEEF 196

QY 181 RIHFTFKFQNNPNPSLKPITQLLGRTHATGIRKVVRELFTNNGARKNAFKILILI 240
Db 197 RIHFTFKFQNNPNPSLKPITQLLGRTHATGIRKVVRELFTNNGARKNAFKILIVI 256

RESULT 5
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-30

Query Match 99.8%; Score 5862; DB 14; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAAQNGSLYQCDYSTGSCBPI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAAQNGSLYQCDYSTGSCBPI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFLFGSNLRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFLFGSNLRQPOK 136

QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEPISTVMEOLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEPISTVMEOLKSKTFLSLMOYSEEF 196

QY 181 RIHFTFKFQNNPNPSLKPITQLLGRTHATGIRKVVRELFTNNGARKNAFKILILI 240
Db 197 RIHFTFKFQNNPNPSLKPITQLLGRTHATGIRKVVRELFTNNGARKNAFKILIVI 256

APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-3

Query Match 99.8%; Score 5862; DB 10; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAAQNGSLYQCDYSTGSCBPI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAAQNGSLYQCDYSTGSCBPI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFLFGSNLRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFLFGSNLRQPOK 136

QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEPISTVMEOLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEPISTVMEOLKSKTFLSLMOYSEEF 196

QY 181 RIHFTFKFQNNPNPSLKPITQLLGRTHATGIRKVVRELFTNNGARKNAFKILILI 240
Db 197 RIHFTFKFQNNPNPSLKPITQLLGRTHATGIRKVVRELFTNNGARKNAFKILIVI 256

QY 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPFSEKSRQBLNTVASKPRDHVQIN 300
Db 257 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPFSEKSRQBLNTVASKPRDHVQIN 316

QY 301 NFEALKTIONLREKIPAIETGTGSSSSSEHMSQSGFSAATNSGPLLSTVGSYDMAG 360
Db 317 NFEALKTIONLREKIPAIETGTGSSSSSEHMSQSGFSAATNSGPLLSTVGSYDMAG 376

QY 361 GVFLYTSKSKSTFTNMTVRVDSMDMDAYLGAAAILLNRVQSLVGLGAPRYCHIGLVAMFR 420
Db 377 GVFLYTSKSKSTFTNMTVRVDSMDMDAYLGAAAILLNRVQSLVGLGAPRYCHIGLVAMFR 436

QY 421 QNTGHWESNANVKQTQI GAYFGASLCSVDVDSNGSTDLVLIGAPHYVEOTRGQVSCPL 480
Db 437 QNTGHWESNANVKQTQI GAYFGASLCSVDVDSNGSTDLVLIGAPHYVEOTRGQVSCPL 496

QY 481 PRGORARWQCDALVYGEQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYTLF 540
Db 497 PRGORARWQCDALVYGEQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYTLF 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLSGQDLTWDGLVDLTGAGQHVILLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLSGQDLTWDGLVDLTGAGQHVILLRSQ 616

QY 601 PVLKVKALMEPNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
Db 617 PVLKVKALMEPNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676

QY 661 YDLALDSGRPHSRAVFNSTKSTRQTOVLGLTQTCFTLKLQPLNCIEDPVPSPVILRLNF 720
Db 677 YDLALDSGRPHSRAVFNSTKSTRQTOVLGLTQTCFTLKLQPLNCIEDPVPSPVILRLNF 736

241 TDGEKFGDPLGYEDVTPADREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVQIN 300
Db |||||
257 TDGEKFGDPLGYEDVTPADREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVQIN 316
Qy |||||
301 NFALXATIQNLREKIFAIEGTOTGSSSFEHMSQEGFSAATISNGPLLSVGSVDWAG 360
Db |||||
317 NFALXATIQNLREKIFAIEGTOTGSSSFEHMSQEGFSAATISNGPLLSVGSVDWAG 376
Qy |||||
361 GVFLYTSKEKSTPINNTRVDSMDNDAYLGVAALIIIRNRVQSLVLGAPRYQHIGLVAMPR 420
Db |||||
377 GVFLYTSKEKSTPINNTRVDSMDNDAYLGVAALIIIRNRVQSLVLGAPRYQHIGLVAMPR 436
Qy |||||
421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLI GAPHYEOTRGQVSVCP 480
Db |||||
437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLI GAPHYEOTRGQVSVCP 496
Qy |||||
481 PRGORARWQCDVLYGEGQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db |||||
497 PRGORARWQCDVLYGEGQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy |||||
541 HGTSGSGISPSHSORITAGSKLSPRLQYFGOSISGGQDLTMDGLVDLTGAGQHVLLRSQ 600
Db |||||
557 HGTSGSGISPSHSORITAGSKLSPRLQYFGOSISGGQDLTMDGLVDLTGAGQHVLLRSQ 616
Qy |||||
601 PVLRVKALMEFNPREVARNVFECDQVWKGKEAGEVRVCLVHVKSTRDLRREGIQSVVT 660
Db |||||
617 PVLRVKALMEFNPREVARNVFECDQVWKGKEAGEVRVCLVHVKSTRDLRREGIQSVVT 676
Qy |||||
661 YDLALDSGRPHSRAVFNETHKSTRTOTVGLTQTCETLKLQPNCEIDPVSPIVLRNPF 720
Db |||||
677 YDLALDSGRPHSRAVFNETHKSTRTOTVGLTQTCETLKLQPNCEIDPVSPIVLRNPF 736
Qy |||||
721 SLVGTPLSAFNGLRPVLAEQAQLFTALPPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db |||||
737 SLVGTPLSAFNGLRPVLAEQAQLFTALPPFEKNCNDNICQDDLSITFSFMSLDCLVVG 796
Qy |||||
781 GPREFNVTVVRNDESDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db |||||
797 GPREFNVTVVRNDESDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
Qy |||||
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
Db |||||
857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916
Qy |||||
901 QLELPVKYAVYVMTSHGVSTKYLNFNTASENTSRVMOHQYQVSNLQBSLPSLVFLVPV 960
Db |||||

RESULT 6
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176

Query Match 99.8%; Score 5862; DB 14; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPQBEIVAAANQORSLVOCYDSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPQBEIVAAANQORSLVOCYDSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHGTCSNTYVVKGLCLFPGSLNLRQOQPK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHGTCSNTYVVKGLCLFPGSLNLRQOQPK 136
Qy 121 PPEALRGCPQEDSDIAFLVDGSGSII PHDPRAKEFIITVMEQKSKTLPFLSMOYSSEF 180
Db 137 PPEALRGCPQEDSDIAFLVDGSGSII PHDPRAKEFIITVMEQKSKTLPFLSMOYSSEF 196
Qy 181 RIHFTFKFQNNPNPRSLIKPITQLLGRTHATGIRKVVRELFTNGARKNAFKILILI 240
Db 197 RIHFTFKFQNNPNPRSLIKPITQLLGRTHATGIRKVVRELFTNGARKNAFKILIVI 256
Qy 241 TDGEKFGDPLGYEDVTPADREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVQIN 300
Db 257 TDGEKFGDPLGYEDVTPADREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVQIN 316
Qy 301 NFEALKTIQNLREKIFAIEGTOTGSSSFEHMSQEGFSAATISNGPLLSVGSVDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTOTGSSSFEHMSQEGFSAATISNGPLLSVGSVDWAG 376
Qy 361 GVFLYTSKEKSTPINNTRVDSMDNDAYLGVAALIIIRNRVQSLVLGAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTPINNTRVDSMDNDAYLGVAALIIIRNRVQSLVLGAPRYQHIGLVAMPR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLI GAPHYEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLI GAPHYEOTRGQVSVCP 496
Qy 481 PRGORARWQCDVLYGEGQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGORARWQCDVLYGEGQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSORITAGSKLSPRLQYFGOSISGGQDLTMDGLVDLTGAGQHVLLRSQ 600
Db 557 HGTSGSGISPSHSORITAGSKLSPRLQYFGOSISGGQDLTMDGLVDLTGAGQHVLLRSQ 616
Qy 601 PVLRVKALMEFNPREVARNVFECDQVWKGKEAGEVRVCLVHVKSTRDLRREGIQSVVT 660
Db 617 PVLRVKALMEFNPREVARNVFECDQVWKGKEAGEVRVCLVHVKSTRDLRREGIQSVVT 676
Qy 661 YDLALDSGRPHSRAVFNETHKSTRTOTVGLTQTCETLKLQPNCEIDPVSPIVLRNPF 720
Db 677 YDLALDSGRPHSRAVFNETHKSTRTOTVGLTQTCETLKLQPNCEIDPVSPIVLRNPF 736
Qy 721 SLVGTPLSAFNGLRPVLAEQAQLFTALPPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNGLRPVLAEQAQLFTALPPFEKNCNDNICQDDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDESDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDESDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFNTASENTSRVMOHQYQVSNLQBSLPSLVFLVPV 960
Db |||||

Db 917 QLELPVKAVYVWVTSHGVSTKYLNTASNTSRVMOHQYQVSNLQORSLSPLSLVFLVPV 976
Qy 961 RLNTQVIMDRPQVTFSENLSTCHTKERLPSSDPLAELRKAPVNVNCSIAVCQRIQCDDIP 1020
Db 977 RLNTQVIMDRPQVTFSENLSTCHTKERLPSSDPLAELRKAPVNVNCSIAVCQRIQCDDIP 1036
Qy 1021 PFGIOEEFNATLKGNSLSPDWYIKTSHNHLIIVSTAEILLFNDVFTLLPQGGAFVRSQTEP 1080
Db 1037 PFGIOEEFNATLKGNSLSPDWYIKTSHNHLIIVSTAEILLFNDVFTLLPQGGAFVRSQTEP 1096
Qy 1081 KVEPEVEPNPLPLIVGSSVGGILLALITAAALYKLGFPKQYKQVNMWSEGGPPGABPQ 1137
Db 1097 KVEPEVEPNPLPLIVGSSVGGILLALITAAALYKLGFPKQYKQVNMWSEGGPPGABPQ 1153

RESULT 7
US-09-902-481A-6
; Sequence 6, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimooka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-6

Query Match 99.6%; Score 5855; DB 10; Length 1137;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1127; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQNRAGFGQSVVQLQGSRRVVGAPQEIIVAAHQRSGLYQCDYSTGSCBEI 60
Db 1 FNLDTENAMTFQNRAGFGQSVVQLQGSRRVVGAPQEIIVAAHQRSGLYQCDYSTGSCBEI 60
Qy 61 RLQVPEAVNMSGLSLAATTSPPQLACGPTVHQTCSENTVYKGLCFLFGSNLRQOPQK 120
Db 61 RLQVPEAVNMSGLSLAATTSPPQLACGPTVHQTCSENTVYKGLCFLFGSNLRQOPQK 120
Qy 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDPFRAXEFISTVMEQLKSKTFLFSLMOYSEEP 180
Db 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDPFRAXEFISTVMEQLKSKTFLFSLMOYSEEP 180
Qy 181 RIHFTKSPQNNPRLSLKPIITOLGRTHATGIRKVVRELVNITNGARKNAFKILLI 240
Db 181 RIHFTKSPQNNPRLSLKPIITOLGRTHATGIRKVVRELVNITNGARKNAFKILLI 240
Qy 241 TDGEKFGDPLGYEDVPEADREGVIRVIGVGDAFRSEKSRQELNTVASKPPEDHVFQIN 300
Db 241 TDGEKFGDPLGYEDVPEADREGVIRVIGVGDAFRSEKSRQELNTVASKPPEDHVFQIN 300
Qy 301 NFPAKTIQNLREKPIAIGTQTSSESSFEHMSQEGFSAATISNGPILSTVGSYDMAG 360
Db 301 NFPAKTIQNLREKPIAIGTQTSSESSFEHMSQEGFSAATISNGPILSTVGSYDMAG 360
Qy 361 GVFLYTSKEKSTFNTMTVDSMDNDAYLGVAAILLNRRVQSLVGLAPRYQHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFNTMTVDSMDNDAYLGVAAILLNRRVQSLVGLAPRYQHIGLVAMFR 420
Qy 421 QNTGMESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEQTRGGQSVCPCL 480

Db 421 QNTGMESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEQTRGGQSVCPCL 480
Qy 481 PRGORARWQCDALVYGBQGPWGRFGAALTIVLDVNGDKLTDVAIGAPGEDNEGAVYLF 540
Db 481 PRGORARWQCDALVYGBQGPWGRFGAALTIVLDVNGDKLTDVAIGAPGEDNEGAVYLF 540
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTGAGQGHVLLLRSQ 600
Db 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTGAGQGHVLLLRSQ 600
Qy 601 PVLRVKAIMFENPREVARNVFECDQVWVGKEAGEVRVCLHVQKSTDRRLREGQIQSVVT 660
Db 601 PVLRVKAIMFENPREVARNVFECDQVWVGKEAGEVRVCLHVQKSTDRRLREGQIQSVVT 660
Qy 661 YDLALDSGRPHSRAVFNETKNSRROTQVLGLTCTETLKLQLPNCIEDPVSPIVLRINF 720
Db 661 YDLALDSGRPHSRAVFNETKNSRROTQVLGLTCTETLKLQLPNCIEDPVSPIVLRINF 720
Qy 721 SLVGTPLSAFGLNRPVLAEDRQRLFTALPPPKKCGNDNICQDDLSITFSFMSLDCLVVG 780
Db 721 SLVGTPLSAFGLNRPVLAEDRQRLFTALPPPKKCGNDNICQDDLSITFSFMSLDCLVVG 780
Qy 781 GPRESFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQORSWRLACESASSTEV 840
Db 781 GPRESFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQORSWRLACESASSTEV 840
Qy 841 SGALKSTSCSINHPIFPENSESVTNITFDVDSKASLGNKLLKANVTSENKXPRNTKTEP 900
Db 841 SGALKSTSCSINHPIFPENSESVTNITFDVDSKASLGNKLLKANVTSENKXPRNTKTEP 900
Qy 901 QLELPVKAVYVWVTSHGVSTKYLNTASNTSRVMOHQYQVSNLQORSLSPLSLVFLVPV 960
Db 901 QLELPVKAVYVWVTSHGVSTKYLNTASNTSRVMOHQYQVSNLQORSLSPLSLVFLVPV 960
Qy 961 RLNTQVIMDRPQVTFSENLSTCHTKERLPSSDPLAELRKAPVNVNCSIAVCQRIQCDDIP 1020
Db 961 RLNTQVIMDRPQVTFSENLSTCHTKERLPSSDPLAELRKAPVNVNCSIAVCQRIQCDDIP 1020
Qy 1021 PFGIOEEFNATLKGNSLSPDWYIKTSHNHLIIVSTAEILLFNDVFTLLPQGGAFVRSQTEP 1080
Db 1021 PFGIOEEFNATLKGNSLSPDWYIKTSHNHLIIVSTAEILLFNDVFTLLPQGGAFVRSQTEP 1080
Qy 1081 KVEPEVEPNPLPLIVGSSVGGILLALITAAALYKLGFPKQYKQVNMWSEGGPPGABPQ 1137
Db 1081 KVEPEVEPNPLPLIVGSSVGGILLALITAAALYKLGFPKQYKQVNMWSEGGPPGABPQ 1137

RESULT 8
US-09-902-481A-4
; Sequence 4, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimooka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-4

Query Match 99.6%; Score 5851; DB 10; Length 1137;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGQSVVOLQSGRVVVGAPQBIIVAAHQSGSLYQCDYSTGSCPEI 60
Db 1 FNLDTENAMTFOENARGFGQSVVOLQSGRVVVGAPQBIIVAAHQSGSLYQCDYSTGSCPEI 60

Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120
Db 61 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120

Qy 121 FPEALRGCPQSDSDIAFLVDSGSIIPHDPRAKEFTSTVMEQLKSKTLFSLMQYSEEF 180
Db 121 FPEALRGCPQSDSDIAFLVDSGSIIPHDPRAKEFTSTVMEQLKSKTLFSLMQYSEEF 180

Qy 181 RIHFTFKFQNNPNSRLIKPIITOLLGRTHATGIRKVVRELFTNGARKNAFKILILI 240
Db 181 RIHFTFKFQNNPNSRLIKPIITOLLGRTHATGIRKVVRELFTNGARKNAFKILILI 240

Qy 241 TDGKFGDPLGYEDVIPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVQIN 300
Db 241 TDGKFGDPLGYEDVIPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVQIN 300

Qy 301 NFEALKTIONLRKIPAIETGOTGSSSSFEHMSQSGFSAITNSGPDLLSTVGSYDMAG 360
Db 301 NFEALKTIONLRKIPAIETGOTGSSSSFEHMSQSGFSAITNSGPDLLSTVGSYDMAG 360

Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420

Qy 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQSVCPIL 480
Db 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQSVCPIL 480

Qy 481 PRGQARWQCDVLYGEGQGPWKEFGAALTIVLGVNDGDKLTDVAIGAPGEDNEGAVYLF 540
Db 481 PRGQARWQCDVLYGEGQGPWKEFGAALTIVLGVNDGDKLTDVAIGAPGEDNEGAVYLF 540

Qy 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTAVGAGHYLLRSQ 600
Db 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTAVGAGHYLLRSQ 600

Qy 601 PVLAVKALMEPNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVYT 660
Db 601 PVLAVKALMEPNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVYT 660

Qy 661 YDLALDSGRPHSRVAFNETKNSRTRQVGLTQTCETLKLQLENCIEDPVSPIVLRNF 720
Db 661 YDLALDSGRPHSRVAFNETKNSRTRQVGLTQTCETLKLQLENCIEDPVSPIVLRNF 720

Qy 721 SLVGTPLSARCNLAPVLAEDAQRLPTALFPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 721 SLVGTPLSARCNLAPVLAEDAQRLPTALFPEKNCNDNICODDLSITFSFMSLDCLVVG 780

Qy 781 GPREFNVTTVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 781 GPREFNVTTVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840

Qy 841 SGALKSTSCSINHPIPEPNSVNTNITFDVDSKASLGNKLLIKANVTSENNMPRTNKTET 900
Db 841 SGALKSTSCSINHPIPEPNSVNTNITFDVDSKASLGNKLLIKANVTSENNMPRTNKTET 900

Qy 901 QLELPVKYAVTMVTSYHGVSTKYLNFTASENTSRVMQHQVQSNLQORSILPISLVLFPV 960
Db 901 QLELPVKYAVTMVTSYHGVSTKYLNFTASENTSRVMQHQVQSNLQORSILPISLVLFPV 960

Qy 961 RLNQTIVDRQVTFSENLSTCHTKERLPSHSDFLAELKAPVWNCIAVCQRIQCDIP 1020
Db 961 RLNQTIVDRQVTFSENLSTCHTKERLPSHSDFLAELKAPVWNCIAVCQRIQCDIP 1020

Qy 1021 FFGIQEFNATLKGNSLSPDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQCAFVRSQTET 1080

Db 1021 FFGIQEFNATLKGNSLSPDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQCAFVRSQTET 1080

Qy 1081 KVEPFEPVNPPLIVGVSSVGGILLALLIITAAALVKLGFPRKQYKDMSEGPPGAEPO 1137
Db 1081 KVEPFEPVNPPLIVGVSSVGGILLALLIITAAALVKLGFPRKQYKDMSEGPPGAEPO 1137

RESULT 9
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; TITLE OF INVENTION: DESIRED CONFORMATION AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: CHN-002CP
; CURRENT APPLICATION NUMBER: US/09/945,265
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,700
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-265-4

Query Match 99.5%; Score 5846.5; DB 9; Length 1152;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 1; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFOENARGFGQSVVOLQSGRVVVGAPQBIIVAAHQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVOLQSGRVVVGAPQBIIVAAHQSGSLYQCDYSTGSCPEI 76

Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 136

Qy 121 FPEALRGCPQSDSDIAFLVDSGSIIPHDPRAKEFTSTVMEQLKSKTLFSLMQYSEEF 180
Db 137 FPEALRGCPQSDSDIAFLVDSGSIIPHDPRAKEFTSTVMEQLKSKTLFSLMQYSEEF 196

Qy 181 RIHFTFKFQNNPNSRLIKPIITOLLGRTHATGIRKVVRELFTNGARKNAFKILILI 240
Db 197 RIHFTFKFQNNPNSRLIKPIITOLLGRTHATGIRKVVRELFTNGARKNAFKILIVI 256

Qy 241 TDGKFGDPLGYEDVIPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVQIN 300
Db 257 TDGKFGDPLGYEDVIPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVQIN 316

Qy 301 NFEALKTIONLRKIPAIETGOTGSSSSFEHMSQSGFSAITNSGPDLLSTVGSYDMAG 360
Db 317 NFEALKTIONLRKIPAIETGOTGSSSSFEHMSQSGFSAITNSGPDLLSTVGSYDMAG 376

Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436

Qy 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQSVCPIL 480
Db 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQSVCPIL 496

Qy 481 PRGQARWQCDVLYGEGQGPWKEFGAALTIVLGVNDGDKLTDVAIGAPGEDNEGAVYLF 540
Db 497 PRGQARWQCDVLYGEGQGPWKEFGAALTIVLGVNDGDKLTDVAIGAPGEDNEGAVYLF 555

Qy 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTAVGAGHYLLRSQ 600

Db 556 HGTSGSISPSHSQRIAGSKLSPLQYFGQSGLQDGLTMDGLVDLTGAGQHVLLRSQ 615
QY 601 PVLRYKALMEFNPREFVARNFECDQVVKGEAGEVRVCLHVKQSTRDLREGQIQSVVT 660
Db 616 PVLRYKALMEFNPREFVARNFECDQVVKGEAGEVRVCLHVKQSTRDLREGQIQSVVT 675
QY 661 YDLALDSGRPHSAFVAFNETKSTERTQVGLGLTQTCETLKLQLPNCIEDPVSPIVLRANF 720
Db 676 YDLALDSGRPHSAFVAFNETKSTERTQVGLGLTQTCETLKLQLPNCIEDPVSPIVLRANF 735
QY 721 SLVGTPLSAFGLNLRPLVLAEDAQRLLFTALFFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFGLNLRPLVLAEDAQRLLFTALFFPEKNCNDNICQDDLSITFSFMSLDCLVVG 795
QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLVSKYKUSLONORSORSWRLACESASTEV 840
Db 796 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLVSKYKUSLONORSORSWRLACESASTEV 855
QY 841 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASIGNKLLKANVTSENNMPTNKTEF 900
Db 856 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASIGNKLLKANVTSENNMPTNKTEF 915
QY 901 QLELPVKYAVVMVYTSHGVSQTKYLNFTASENTRSVMOHQYQVSNIGORSILPSLPLVPV 960
Db 916 QLELPVKYAVVMVYTSHGVSQTKYLNFTASENTRSVMOHQYQVSNIGORSILPSLPLVPV 975
QY 961 RLNQTVIWDROPQVTFENSELSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRIQCDIP 1020
Db 976 RLNQTVIWDROPQVTFENSELSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRIQCDIP 1035
QY 1021 PFGIQEBFNATLKNLSFDWYIKTSHNHLLIVSTAEILFNDVSFTLLPQCGAFVRSQTEF 1080
Db 1036 PFGIQEBFNATLKNLSFDWYIKTSHNHLLIVSTAEILFNDVSFTLLPQCGAFVRSQTEF 1095
QY 1081 KVEPFEPVNPPLIVGSSVGGLLLLALITAAALYKLGFFPKRYQKDMMSBEGPPGABPQ 1137
Db 1096 KVEPFEPVNPPLIVGSSVGGLLLLALITAAALYKLGFFPKRYQKDMMSBEGPPGABPQ 1152

RESULT 10

US-09-902-481A-3

; Sequence 3, Application US/09902481A

; Publication No. US20030054440A1

; GENERAL INFORMATION:

; APPLICANT: Springer, Timothy

; APPLICANT: Shimooka, Motomu

; APPLICANT: Shifman, Julia

; APPLICANT: Mayo, Stephen

; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY

; FILE REFERENCE: A-70586-1/RT/RMS/RMK

; CURRENT APPLICATION NUMBER: US/09/502,481A

; CURRENT FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/216,600

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 1137

; TYPE: PRT

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: synthetic

US-09-902-481A-3

Query Match

Best Local Similarity 99.3%; Score 5836; DB 10; Length 1137;

Matches 1124; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAHQSLVQCDYSTGSCPEI 60
Db 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAHQSLVQCDYSTGSCPEI 60
QY 61 RLQVPEVAVNMSLGLSLAATTPPQLLACGPTVHQTCSNTYVVKGLCFLEGSNLRQQPQK 120

RESULT 11

Db 61 RLQVPEVAVNMSLGLSLAATTPPQLLACGPTVHQTCSNTYVVKGLCFLEGSNLRQQPQK 120
QY 121 FPRALRGCPQSDSIAFLVDGSGSIIIPHDERRAKEPISTVMBOLKKSKTLFSLMXYSEBF 180
Db 121 FPRALRGCPQSDSIAFLVDGSGSIIIPHDERRAKEPISTVMBOLKKSKTLFSLMXYSEBF 180
QY 181 RIHTTFKQFNPNRSLIKPIITOLLORTHTATGIRKQVRELNTNNGARKNAPKILILI 240
Db 181 RIHTTFKQFNPNRSLIKPIITOLLORTHTATGIRKQVRELNTNNGARKNAPKILILI 240
QY 241 TDGKFGDPLGYEDVIVPEADREGVIRVYVIGVDAFRSEKSRQELNVTASKPPDRHVFQIN 300
Db 241 TDGKFGDPLGYEDVIVPEADREGVIRVYVIGVDAFRSEKSRQELNVTASKPPDRHVFQIN 300
QY 301 NFRALKTIQVLRKIKPIAIEGTQTGSSSSPEHMSQSGPSAAITNSKGPLLSTVGSVDWAG 360
Db 301 NFRALKTIQVLRKIKPIAIEGTQTGSSSSPEHMSQSGPSAAITNSKGPLLSTVGSVDWAG 360
QY 361 GVPLYTSKESKSTFINMTRVDSMDNDAYLGVAALILNRVQSLVLCAPRQHIHGLVAMFR 420
Db 361 GVPLYTSKESKSTFINMTRVDSMDNDAYLGVAALILNRVQSLVLCAPRQHIHGLVAMFR 420
QY 421 ONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVLIIGAPHYETGREGQVSCVL 480
Db 421 ONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVLIIGAPHYETGREGQVSCVL 480
QY 481 PRGQARWQCDVAVLYGEGQGFPGCFGAALTIVLGDVANGDKLTDVAIGAPGEEDNRGAVILF 540
Db 481 PRGQARWQCDVAVLYGEGQGFPGCFGAALTIVLGDVANGDKLTDVAIGAPGEEDNRGAVILF 540
QY 541 HGTSGSISPSHSQRIAGSKLSPLQYFGQSGLQDGLTMDGLVDLTGAGQHVLLRSQ 600
Db 541 HGTSGSISPSHSQRIAGSKLSPLQYFGQSGLQDGLTMDGLVDLTGAGQHVLLRSQ 600
QY 601 PVLRYKALMEFNPREFVARNFECDQVVKGEAGEVRVCLHVKQSTRDLREGQIQSVVT 660
Db 601 PVLRYKALMEFNPREFVARNFECDQVVKGEAGEVRVCLHVKQSTRDLREGQIQSVVT 660
QY 661 YDLALDSGRPHSAFVAFNETKSTERTQVGLGLTQTCETLKLQLPNCIEDPVSPIVLRANF 720
Db 661 YDLALDSGRPHSAFVAFNETKSTERTQVGLGLTQTCETLKLQLPNCIEDPVSPIVLRANF 720
QY 721 SLVGTPLSAFGLNLRPLVLAEDAQRLLFTALFFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 721 SLVGTPLSAFGLNLRPLVLAEDAQRLLFTALFFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLVSKYKUSLONORSORSWRLACESASTEV 840
Db 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLVSKYKUSLONORSORSWRLACESASTEV 840
QY 841 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASIGNKLLKANVTSENNMPTNKTEF 900
Db 841 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASIGNKLLKANVTSENNMPTNKTEF 900
QY 901 QLELPVKYAVVMVYTSHGVSQTKYLNFTASENTRSVMOHQYQVSNIGORSILPSLPLVPV 960
Db 901 QLELPVKYAVVMVYTSHGVSQTKYLNFTASENTRSVMOHQYQVSNIGORSILPSLPLVPV 960
QY 961 RLNQTVIWDROPQVTFENSELSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRIQCDIP 1020
Db 961 RLNQTVIWDROPQVTFENSELSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRIQCDIP 1020
QY 1021 PFGIQEBFNATLKNLSFDWYIKTSHNHLLIVSTAEILFNDVSFTLLPQCGAFVRSQTEF 1080
Db 1021 PFGIQEBFNATLKNLSFDWYIKTSHNHLLIVSTAEILFNDVSFTLLPQCGAFVRSQTEF 1080
QY 1081 KVEPFEPVNPPLIVGSSVGGLLLLALITAAALYKLGFFPKRYQKDMMSBEGPPGABPQ 1137
Db 1081 KVEPFEPVNPPLIVGSSVGGLLLLALITAAALYKLGFFPKRYQKDMMSBEGPPGABPQ 1137

US-10-116-275-204
 ; Sequence 204, Application US/10116275
 ; Publication No. US20030211476A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eian Pharmaceutical Technology
 ; APPLICANT: O'Mahony, Daniel J.
 ; APPLICANT: Brayden, David
 ; APPLICANT: Byrne, Daragh
 ; APPLICANT: Lambkin, Imelda
 ; APPLICANT: Higgins, Lisa
 ; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
 ; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
 ; FILE REFERENCE: E1067/20087
 ; CURRENT APPLICATION NUMBER: US/10/116,275
 ; CURRENT FILING DATE: 2002-10-04
 ; NUMBER OF SEQ ID NOS: 349
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 204
 ; LENGTH: 1163
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-116-275-204

Query Match 59.0%; Score 3469; DB 15; Length 1163;
 Best Local Similarity 61.0%; Pred. No. 2.3e-309;
 Matches 689; Conservative 141; Mismatches 293; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFORNARGFQSGVQVLOGSVVVGAPQEIIVANQORSLVOCYDSTGSCPEI 60
 DB 20 FNLDTENAMTFORNARGFQSGVQVLOGSVVVGAPQEIIVANQORSLVOCYDSTGSCPEI 79

QY 61 RLQVPEAVNMVSLGSLAATTSPQOLACGPTVHQTCSNTVYVKGCLFPGSLNLRQPOK 120
 DB 80 GLQVPEAVNMVSLGSLAATTSPQOLACGPTVHQTCSNTVYVKGCLFPGSLNLRQPOK 137

QY 121 FPEALRGCEQEDSDIAFLVDGSGSIIPHDPRRAKFFISTVMEQKKSKTLFSLMQLVSEF 180
 DB 138 LPVSRQECPRQEQDIFVLIDGSGSISSRNFAFMNFAVVISQFORPSTQFSLMQFSNKF 197

QY 181 RHFTFKQNNPNSLSKIPITOLLGRTHGTGIRKVVRELFNTNGARKNAKILILI 240
 DB 198 QTHFTFEEFRSTNPLSLASVHQLQGGFTYATQVVRHVFHASYGARRDATKILIVI 257

QY 241 TDCEKFGDPLGYEDVPEADREGVIRVIGVGDFAFSEKSRQELNATVASKPPRDRHVQIN 300
 DB 258 TDCKGSDSLDKVDPMDAAGIIRYAGVGLAFQNRNSWKLNDIAKSPQSEHIFKVE 317

QY 301 NFALKTIONOLREKIPALEGTOTGSSSFEHMSQEGFSAALTSNGPLISTVGSVDWAG 360
 DB 318 DFDALDKIQNLKEKIFAIEGTETSSSFEHMSQEGFSAALTSNGPLISTVGSVDWAG 377

QY 361 GVLPLYSKESKSTFINTNRVDSMDNDAYLGAAAIILNRNVQSILVLCAPRYQHIGLVANFR 420
 DB 378 GATLYPNKSPPTINNSQENVDNRDYSYVSTELAKWGVQSILVLCAPRYQHIGLVANFR 437

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETRGQVSVCPJ 480
 DB 438 QVSRQWRMAEVTGTQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYETRGQVSVCPJ 497

QY 481 PRQBARWODAVLYGEQOPWGRFGAALTVLGDVNGDKLTDVAICAPGEEDNRGAVILF 540
 DB 498 PRGWR-RWMCDAVLYGEQOPWGRFGAALTVLGDVNGDKLTDVAICAPGEEDNRGAVILF 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPRLAYFGQSLSGGQDLTDGGLVDTVGAQGHVLLRSQ 600
 DB 557 HGVLPISGISPSHSQRIAGSKLSPRLAYFGQSLSGGQDLTDGGLVDTVGAQGHVLLRSQ 616

QY 601 PVLKVAIMEFNPREVARNVFECDQVVKKEAGEVRVCLHVOKSTFDRLEQIQSVVT 660
 DB 617 PVLVWVGSMOFIPIABIPRPAFECREQVVSQTLVQSNICLIDIRSKNLLGSDIASSVT 676

QY 661 YDLALDSEGRPHSRVAFNEFNKSTNRQTVLGLTCTCTETLKLQLPNCIEDPVPVILRLNF 720

Db 677 LDALDPEGLSPRATQETKNSRSLRVVLGLKAHCENFNLLPSCVEDSVPTITRLNF 736
 QY 721 SLVGTPLSAFGNLRPLVLAEDAQRLFTALPFFERNCGNDNICODLSITTFPMSLDCLVVG 780
 Db 737 TLVGKPLAFNLRPLMLAQAQRYFTASLPEFKNCGADHICQDNLIGISFPGKSLVVG 796
 QY 781 GPREFNVTVVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
 Db 797 SNLELAEVAVVNDGDSYGTTFTHSPAGLSYRYVAGQKQGLRSLHLTCDSPVGV- 854
 QY 841 SGALKSTCSINHPFENSEEVNITTFDVSASIKASGNKLLKANVTSENMMPTNKTEF 900
 Db 855 SQGTWSTSCRNHLIFRGAQITPLATEDVSPRAVLGDRLLLTANVSENNTPTSTKTF 914

QY 901 QLELPVKYAVVMTVSGVSTKYLNPAS-ENTSRVMOHOYOVSNLQORSLPISLVFLVP 959
 Db 915 QLELPVKYAVVMTVSGVSTKYLNPAS-ENTSRVMOHOYOVSNLQORSLPISLVFLVP 974

QY 960 VRLNQTVIWDPRQVTFSENLSSTCHTERLPFSDPLAEIRKAPVAVNCSIAVOCRIQCDI 1019
 Db 975 VELNQEAVMVDVESHQNPCLRCSEKIPASDPLAHIQKNEVPLDCSIAGCLFRCDV 1034

QY 1020 PFGIQIEFNATLKNLSFDWYIKTSHNLLIUVSTAILNDVSTLPGOGAFVRSQTE 1079
 Db 1035 PSFSVQSELDFTLKNISFGWVRQILOKCVSVSVAEITPDTSVISQLPQGEAFNRAGTT 1094

QY 1080 TKVPPFEVENPLIVGSSVGLLILALITAAALYKLAGFFKQYKDMKSE 1128
 Db 1095 TVLEKXVHNPPLIVGSSIGLILALITAVLYKVGFFKQYKEMKSE 1143

RESULT 12
 US-09-350-259-4
 ; Sequence 4, Application US/09350259
 ; Patent No. US20020062008A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gallatin, Michael W.
 ; APPLICANT: Van der Vieren, Monica
 ; TITLE OF INVENTION: NO. US20020062008A1el Human 2
 ; FILE REFERENCE: 27866/35004
 ; CURRENT APPLICATION NUMBER: US/09/350,259
 ; CURRENT FILING DATE: 1999-07-08
 ; EARLIER APPLICATION NUMBER: 09/193,043
 ; EARLIER FILING DATE: 1998-11-16
 ; EARLIER APPLICATION NUMBER: 08/173,497
 ; EARLIER FILING DATE: 1993-12-23
 ; EARLIER APPLICATION NUMBER: 08/286,889
 ; EARLIER FILING DATE: 1994-08-05
 ; EARLIER APPLICATION NUMBER: 08/362,652
 ; EARLIER FILING DATE: 1994-12-21
 ; EARLIER APPLICATION NUMBER: 08/943,363
 ; EARLIER FILING DATE: 1997-10-03
 ; NUMBER OF SEQ ID NOS: 114
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 1163
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-350-259-4

Query Match 58.6%; Score 3446; DB 9; Length 1163;
 Best Local Similarity 60.8%; Pred. No. 3e-307;
 Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFORNARGFQSGVQVLOGSVVVGAPQEIIVANQORSLVOCYDSTGSCPEI 60
 DB 20 FNLDTENAMTFORNARGFQSGVQVLOGSVVVGAPQEIIVANQORSLVOCYDSTGSCPEI 79

QY 61 RLQVPEAVNMVSLGSLAATTSPQOLACGPTVHQTCSNTVYVKGCLFPGSLNLRQPOK 120
 Db 80 GLQVPEAVNMVSLGSLAATTSPQOLACGPTVHQTCSNTVYVKGCLFPGSLNLRQPOK 137

QY 121 FPEALRGCEQEDSDIAFLVDGSGSIIPHDPRRAKFFISTVMEQKKSKTLFSLMQLVSEF 180

138 LFVSRQECPRQEQDVLFDLSDGSSLSRNFAATMNFVRAVISQFQRPSTQPSLMQFSNKF 197
181 RIHFTFKFQNNPNRSLIKITQOLLGTHRTATGIRKVVRELFNITNGARKNAFKILILI 240
198 QTHFTFEFRSTNLSLASVHQLQGGFTYTATQNVVHFLFASYGARRDAIKILIVI 257
241 TDGKFGDPLGYEDVPEADREGVIRYVIGDAPRSEKSEQLNTVASKPRDPRVQIN 300
258 TDGKFGDPLGYEDVPEADREGVIRYVIGDAPRSEKSEQLNTVASKPRDPRVQIN 317
301 NFEALKTIONLREKIPALBEGTQTSSTSSFEHMSQEGFSAITNSGPELLSTVGSYDWAG 360
318 DFDALKTIONLREKIPALBEGTQTSSTSSFEHMSQEGFSAITNSGPELLSTVGSYDWAG 377
361 GVFLYTSKEKSTFTINMTVDSDMDAYLGYAAIILNRVQSLVGLGAPRYOHIGKAVIFI 420
378 GAFLYPPNMSPTFINNSQENVDRDLSYLGSTELAMKGVQSLVGLGAPRYOHIGKAVIFI 437
421 QNTGWMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYYQTRGGQVSCPL 480
438 QVSRQWRKAEVIGTQIGSYFGASLCSVDVDSNGSTDLVLGAPHYYQTRGGQVSCPL 497
481 PRGQARWOCDAVLYGEOQPMGRFGAALTIVLDVNGDKLTDVAGCAPGEENRGAVILF 540
498 PRGWR-RWCDVAVLYGEOQPMGRFGAALTIVLDVNGDKLTDVAGCAPGEENRGAVILF 556
541 HCTSGSGISPSHSORLAGSKLPRLOYFGQSLSGQDLTMDGLVLTVAOQHVLILRSQ 600
557 HGVLPSPISPSHSORLAGSKLPRLOYFGQSLSGQDLTMDGLVLTVAOQHVLILRSQ 616
601 PVLRYKAIMFNPREVARNVFPCNDQVVKGEAGVRVCLHVOKSTRDLREGQIQSVVT 660
617 PVLWVGVSQMOFIPAEIPRPAFECEQVSEQTLVQSNICLYIDKRSKNLLSRDLQSSVT 676
661 YDLALDSGRPHSRVAFNETKSTRQTVLGTQTCETLKLQPNCEIDPVSPIVLRNF 720
677 LDALAPGRLSPRAIFQSTKRSLSRVVGLKACENFNLLPSCVEDSVPIILRLNF 736
721 SLVGTPLSAGNLRPLVLAEDAQLFTALEPFFBKNCNDNICDDLSITFSPMSLDCLVVG 780
737 TLVGRKELLAFNLRPLMLAALQRYPTASLPFEKNCAGADHICQDNLGISFSPGLKSLVVG 796
781 GRPEFNVTVVNDGEDSTRTQVTFPPLDLSVRKYSTLQNSORSWELACESASTEV 840
797 SNLELNAEVVNDGSDSYGTTITFSHPAGLSRYVABGQKQSLHLTC--CSAPVG 854
841 SGALKETSCSINHPIPEENSEVTFNITFDVDSKASLGNKLLKANVTSENNPRTNKTEP 900
855 SQGTASTSCSINHPIPEENSEVTFNITFDVDSKASLGNKLLKANVTSENNPRTNKTEP 914
901 QLELPKYAVYVVSSEHQFTKYNLFSESEKESHVAMERYQNNLQORDLPVSNFVWP 959
915 QLELPKYAVYVVSSEHQFTKYNLFSESEKESHVAMERYQNNLQORDLPVSNFVWP 974
960 VRLNQTIVDRPOVTFSENLSSCTHTKRLPSHSDFLAELRKPVNVCSIAVCORIQCHI 1019
975 VELNQAVMVDVSHQPSLSCSEKATAPASDFLAHOKNPLVDCSLACGLRCDV 1034
1020 PFGIOEFENATLKNLSPDWYIKTSHNLLIYSTAEILFNDSVFTLLPQCGAFVRSQTE 1079
1035 PPSVQSEELDFTLKNLSPDWYIKTSHNLLIYSTAEILFNDSVFTLLPQCGAFVRSQTE 1094
1080 TKVEPEFVNPPLIVGSSVGLLALITAAVLKLGPFKPKQYKDWSE 1128
1095 TVLEKYKVNHPILIVGSSVGLLALITAAVLKLGPFKPKQYKDWSE 1143

RESULT 13

US-09-891-943-4

; Sequence 4, Application US/09891943

; Publication No. US2003007728A1

; GENERAL INFORMATION:

APPLICANT: Gallatin, Michael W.
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. US2003007728A1 Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/891,943
PRIOR FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 09/193,043
PRIOR FILING DATE: 1998-11-16
PRIOR APPLICATION NUMBER: 08/286,889
PRIOR FILING DATE: 1994-08-05
PRIOR APPLICATION NUMBER: 08/362,652
PRIOR FILING DATE: 1994-12-21
PRIOR APPLICATION NUMBER: 08/943,363
PRIOR FILING DATE: 1997-10-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 1163
TYPE: PRT
ORGANISM: Homo sapiens
US-09-891-943-4

Query Match 58.6%; Score 3446; DB 10; Length 1163;

Best Local Similarity 60.8%; Pred. No. 3e-307;

Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFQENARVARGFQSVQLOQSRVVVGAPQBEIVAAQNGSLYQCDYSTGSCPEI 60
DB 20 FNLDTENAMTFQENARVARGFQSVQLOQSRVVVGAPQBEIVAAQNGSLYQCDYSTGSCPEI 79
QY 61 RLQVPEAVNMSLGLSLAATTPPOLLACQPTVHOTSENTYVKGCLFCLFSGNLRRQOPQK 120
DB 80 GLQVPPAVNMSLGLSLAATTPPOLLACQPTVHOTSENTYVKGCLFCLFSGNLRRQOPQK 137
QY 121 FPEALRQCPQSDSDIAFLVDGSGSIIIPHDPRRAKEPFTVNEQLKKSKTLPSLMQYSEEP 180
DB 138 LFVSRQECPRQEQDVLFDLSDGSSLSRNFAATMNFVRAVISQFQRPSTQPSLMQFSNKF 197
QY 181 RIHFTFKFQNNPNRSLIKITQOLLGTHRTATGIRKVVRELFNITNGARKNAFKILILI 240
DB 198 QTHFTFEFRSTNLSLASVHQLQGGFTYTATQNVVHFLFASYGARRDAIKILIVI 257
QY 241 TDGKFGDPLGYEDVPEADREGVIRYVIGDAPRSEKSEQLNTVASKPRDPRVQIN 300
DB 258 TDGKFGDPLGYEDVPEADREGVIRYVIGDAPRSEKSEQLNTVASKPRDPRVQIN 317
QY 301 NFEALKTIONLREKIPALBEGTQTSSTSSFEHMSQEGFSAITNSGPELLSTVGSYDWAG 360
DB 318 DFDALKTIONLREKIPALBEGTQTSSTSSFEHMSQEGFSAITNSGPELLSTVGSYDWAG 377
QY 361 GVFLYTSKEKSTFTINMTVDSDMDAYLGYAAIILNRVQSLVGLGAPRYOHIGKAVIFI 420
DB 378 GAFLYPPNMSPTFINNSQENVDRDLSYLGSTELAMKGVQSLVGLGAPRYOHIGKAVIFI 437
QY 421 QNTGWMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYYQTRGGQVSCPL 480
DB 438 QVSRQWRKAEVIGTQIGSYFGASLCSVDVDSNGSTDLVLGAPHYYQTRGGQVSCPL 497
QY 481 PRGQARWOCDAVLYGEOQPMGRFGAALTIVLDVNGDKLTDVAGCAPGEENRGAVILF 540
DB 498 PRGWR-RWCDVAVLYGEOQPMGRFGAALTIVLDVNGDKLTDVAGCAPGEENRGAVILF 556
QY 541 HCTSGSGISPSHSORLAGSKLPRLOYFGQSLSGQDLTMDGLVLTVAOQHVLILRSQ 600
DB 557 HGVLPSPISPSHSORLAGSKLPRLOYFGQSLSGQDLTMDGLVLTVAOQHVLILRSQ 616
QY 601 PVLRYKAIMFNPREVARNVFPCNDQVVKGEAGVRVCLHVOKSTRDLREGQIQSVVT 660
DB 617 PVLWVGVSQMOFIPAEIPRPAFECEQVSEQTLVQSNICLYIDKRSKNLLSRDLQSSVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRQTVLGTQTCETLKLQPNCEIDPVSPIVLRNF 720
DB 677 LDALAPGRLSPRAIFQSTKRSLSRVVGLKACENFNLLPSCVEDSVPIILRLNF 736

721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPFKKNCNDNICODDLSITFTSFMSLDCLVVG 780
734 SLVREPIPSQNLAPVLAAGSODLFTASLPFKKNCQDGLCEGLVTLSPSGQLTLTVG 793
781 GPRFNVTVTVRNDGDSYRTQVTFPPPLDLSYKVKSTLQNRORSWRNLACBSASSTEV 840
794 SSLBLNVITVWAGDSYGFVWSLTYTLAGLHRRVSGAQXQPHQSALRLACETV-PTED 852
841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTET 900
853 EG-LRSSRCSVNHPIFHEGSGNGTFTVTFVSYKATLGDRLMLRASSENKASSKATP 911
901 QLELPVKYAVTVYVSHGVSTKYNP-TASENTSRVWQHOYQVSNLQORSPLSLVFLVP 959
912 QLELPVKYAVTVYVSHGVSTKYNP-TASENTSRVWQHOYQVSNLQORSPLSLVFLVP 971
960 VRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELKPAFVYVNCISIAVCORIQDI 1019
972 VLLNGVAVWVWMEAPSQSL--PCVSEKPPQHSDFLTQISKSPMLDCSIADCLQFRCDV 1029
1020 PFFGIQEEFNATLKNLSFDWYIKTSENHLLIVSTABILFNDVSFTLLPGOGAFVRSOTE 1079
1030 PSFSVQBELDFTLKNLSFGWVRETQKKVLVWVABEITFTSVYSQDPGQAFMRAQNE 1089
1080 TKVZEPFVNPPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMWSE 1128
1090 MYLEDEVYNAIPIIMGSSVGGALLLALITATLYKLGFFKHYKEMLED 1138

Search completed: June 7, 2004, 17:38:51
Job time : 42.8905 secs

QY 721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPFKKNCNDNICODDLSITFTSFMSLDCLVVG 780
DB 734 SLVREPIPSQNLAPVLAAGSODLFTASLPFKKNCQDGLCEGLVTLSPSGQLTLTVG 793
QY 781 GPRFNVTVTVRNDGDSYRTQVTFPPPLDLSYKVKSTLQNRORSWRNLACBSASSTEV 840
DB 794 SSLBLNVITVWAGDSYGFVWSLTYTLAGLHRRVSGAQXQPHQSALRLACETV-PTED 852
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTET 900
DB 853 EG-LRSSRCSVNHPIFHEGSGNGTFTVTFVSYKATLGDRLMLRASSENKASSKATP 911
QY 901 QLELPVKYAVTVYVSHGVSTKYNP-TASENTSRVWQHOYQVSNLQORSPLSLVFLVP 959
DB 912 QLELPVKYAVTVYVSHGVSTKYNP-TASENTSRVWQHOYQVSNLQORSPLSLVFLVP 971
QY 960 VRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELKPAFVYVNCISIAVCORIQDI 1019
DB 972 VLLNGVAVWVWMEAPSQSL--PCVSEKPPQHSDFLTQISKSPMLDCSIADCLQFRCDV 1029
QY 1020 PFFGIQEEFNATLKNLSFDWYIKTSENHLLIVSTABILFNDVSFTLLPGOGAFVRSOTE 1079
DB 1030 PSFSVQBELDFTLKNLSFGWVRETQKKVLVWVABEITFTSVYSQDPGQAFMRAQNE 1089
QY 1080 TKVZEPFVNPPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMWSE 1128
DB 1090 MYLEDEVYNAIPIIMGSSVGGALLLALITATLYKLGFFKHYKEMLED 1138

Search completed: June 7, 2004, 17:38:51
Job time : 42.8905 secs

QY 1 FNLDTENAMTFOENARGFGOSVVOLOGSRVVGAPQEIIVANORGSLYOCIDYSTGSCPEI 60
DB 17 FNLDSVEPTIFQDAGFGOSVVOLOGSRVVGAPQEIIVANORGSLYOCIDYSTGSCPEI 76
QY 61 RLQVPEAVNMSLGLSIAATSPOLLACQPTVHOTCSNTYVKGCLFLFGSNLRQPOK 120
DB 77 PLHIREAVNMSLGLSIAATSPOLLACQPTVHOTCSNTYVKGCLFLFGSNLRQPOK 135
QY 121 FPEALRGCPQEDSIAFLVDGSGSIIIPDPRRABFISTVMEQKSKNTLPSLMQYSEEF 180
DB 136 VPDATPCPQEDSIAFLVDGSGSIIIPDPRRABFISTVMEQKSKNTLPSLMQYSEEF 195
QY 181 RIHFTPEKQFNENPRLIPIQLLGRTHATGIRKVVRELENIINGARKNAFKILILI 240
DB 196 KIHFTFTQFTSPSQSLVDPIVOLKGLTFATGILTVTQLFHKNKARKSAKILIVI 255
QY 241 TCGEKFDGPIGYENVIPEADREBGRVIRYVIGVDAFRSEKSRORLNTVASKPPRDHVFQIN 300
DB 256 TDGQKYNDPLEYSDVIPAQEKAGIIRYVIGVDAFRSEKSRORLNTVASKPPRDHVFQIN 315
QY 301 NFEALNTIQNLREKIPIAIGTQGTGSSSTHEHMSQEGPSAAITSNCPILLSTVGSYDNAG 360
DB 316 NFAALGSIQLOKEKIYAVEGTQSRASSSTQHEMSQEGPSAAITSNCPILLSTVGSYDNAG 375
QY 361 GVELYTSKESKSTRINMTVRVDSMDNDAYLVGAAALILRNVRQSLVIGAPRYOHIGLVAMER 420
DB 376 GAFLYPPNMGPTFLNMGQENVMDSDYLGSTELALAKGQNLVIGAPRYOHIGLVAMER 435
QY 421 QNTGWSNANVKTQIGAYFGASLCSVDVDSNGSDTLVIGAPHYETEOTRGQVSVCP 480
DB 436 QVSRQWKKAEVGTQIGAYFGASLCSVDVDSNGSDTLVIGAPHYETEOTRGQVSVCP 495
QY 481 PRGQARWQCDVLYRGQGPWGRFGAALTVLGDVNGDKLTVAGAPGEEDNRGAVYLF 540
DB 496 PRGQARWQCDVLYRGQGPWGRFGAALTVLGDVNGDKLTVAGAPGEEDNRGAVYLF 555
QY 541 RGTSGSISPSHSQRIAGSKLSPRLQVFGSLSGGQDLTMDGLVDTLVGACGHVLLRSQ 600
DB 556 HGASEGISPSHSQRIAGSKLSPRLQVFGSLSGGQDLTMDGLVDTLVGACGHVLLRSQ 615
QY 601 PYLRVKAIMEPNPREVARNFECDNDVVKGEAGEVRCVCLHVKSTDRDLREGOIQSVWT 660
DB 616 FVLKVGAVMRFSPVEAVKVRCEEKPSALRAGDATVCLTIQKSSLDQL--GDIQSVR 673
QY 661 YDLALDGRPHSRVAFNETKSTRQTVGLTQTCTETKLQLENFIEDPVPVILRLNF 720
DB 674 FDLALDPCGLTSRAIFNETKPTLTKRTKTLGLGHCETKLLLPDCVEDVVPVILRLNF 733

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:07:20 ; Search time 18.3484 Seconds
(without alignments)
3199.127 Million cell updates/sec

Title: US-09-902-481B-5
Perfect score: 5876
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pap.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pap.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pap.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pap.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pap.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5862	99.8	1153	1	US-08-173-497-3
2	5862	99.8	1153	1	US-08-286-889-3
3	5862	99.8	1153	1	US-08-485-618-3
4	5862	99.8	1153	1	US-08-362-652-3
5	5862	99.8	1153	2	US-08-605-672-3
6	5862	99.8	1153	2	US-08-482-293A-3
7	5862	99.8	1153	2	US-08-943-363-3
8	5862	99.8	1153	3	US-09-193-043-3
9	5862	99.8	1153	4	US-09-688-307A-3
10	5862	99.8	1153	4	US-09-350-259-3
11	5831.5	99.2	1152	2	US-08-476-062A-43
12	5831.5	99.2	1152	5	PCT-US96-01314-43
13	5831.5	99.2	1152	6	5424399-2
14	3469	59.0	1163	2	US-08-476-062A-44
15	3469	59.0	1163	5	PCT-US96-01314-44
16	3446	58.6	1163	1	US-08-173-497-4
17	3446	58.6	1163	1	US-08-286-889-4
18	3446	58.6	1163	1	US-08-485-618-4
19	3446	58.6	1163	1	US-08-362-652-4
20	3446	58.6	1163	2	US-08-605-672-4
21	3446	58.6	1163	2	US-08-482-293A-4
22	3446	58.6	1163	2	US-08-943-363-4
23	3446	58.6	1163	3	US-09-193-043-4
24	3446	58.6	1163	4	US-09-688-307A-4
25	3446	58.6	1163	4	US-09-350-259-4
26	3411	58.0	1161	1	US-08-173-497-2
27	3411	58.0	1161	1	US-08-286-889-2

28	3411	58.0	1161	1	US-08-485-618-2	Sequence 2, Appli
29	3411	58.0	1161	1	US-08-362-652-2	Sequence 2, Appli
30	3411	58.0	1161	2	US-08-605-672-2	Sequence 2, Appli
31	3411	58.0	1161	2	US-08-482-293A-2	Sequence 2, Appli
32	3411	58.0	1161	2	US-08-943-363-2	Sequence 2, Appli
33	3411	58.0	1161	3	US-09-193-043-2	Sequence 2, Appli
34	3411	58.0	1161	4	US-09-688-307A-2	Sequence 2, Appli
35	3411	58.0	1161	4	US-09-350-259-2	Sequence 2, Appli
36	3395.5	57.8	1161	1	US-08-485-618-99	Sequence 99, Appl
37	3395.5	57.8	1161	2	US-08-605-672-99	Sequence 99, Appl
38	3395.5	57.8	1161	2	US-08-482-293A-99	Sequence 99, Appl
39	3395.5	57.8	1161	2	US-08-943-363-99	Sequence 99, Appl
40	3395.5	57.8	1161	3	US-09-193-043-99	Sequence 99, Appl
41	3395.5	57.8	1161	4	US-09-688-307A-99	Sequence 99, Appl
42	3395.5	57.8	1161	4	US-09-350-259-99	Sequence 99, Appl
43	3232.5	55.0	1161	3	US-09-193-043-55	Sequence 55, Appl
44	3232.5	55.0	1161	4	US-09-688-307A-55	Sequence 55, Appl
45	3232.5	55.0	1161	4	US-09-350-259-55	Sequence 55, Appl

ALIGNMENTS

RESULT 1
US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; TITLE OF INVENTION: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Garstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-173-497-3

Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFQENARGFGQSVVQLGSRVVGAPQEIIVAAQNGRGLYQCDYTGSGCEPI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLGSRVVGAPQEIIVAAQNGRGLYQCDYTGSGCEPI 76

QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVVKGLCFPLGNSLRQOPQK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVVKGLCFPLGNSLRQOPQK 136
QY 121 PPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTSLFSLMOYSEEF 180
DB 137 PPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTSLFSLMOYSEEF 196
QY 181 RHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFTNGARKNAKILILI 240
DB 197 RHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFTNGARKNAKILIVI 256
QY 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVFQIN 300
DB 257 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVFQIN 316
QY 301 NFEALKTIONLREKIFAIECTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIONLREKIFAIECTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFIINWTRVDSMDNDAYLGAAAILRNRVOSLVLGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFIINWTRVDSMDNDAYLGAAAILRNRVOSLVLGAPRYOHIGLVAMFR 436
QY 421 QNTGMWESNANYKGTQIGAYFGASLCSDVDNSGSTDLVLIGAPHYYEQTRGGQVSVCP 480
DB 437 QNTGMWESNANYKGTQIGAYFGASLCSDVDNSGSTDLVLIGAPHYYEQTRGGQVSVCP 496
QY 481 PRGQARWQCDVLYGEOGPGRFGAALTVDGVNGDKLTDVAIGAPGEDNKGAVLYF 540
DB 497 PRGQARWQCDVLYGEOGPGRFGAALTVDGVNGDKLTDVAIGAPGEDNKGAVLYF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTVAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTVAQGHVLLRSQ 616
QY 601 PVLKVAIMEFPREVARVFCNDQVKGKAGEVRVCLHVOKSTRDLREGQIQSVVT 660
DB 617 PVLKVAIMEFPREVARVFCNDQVKGKAGEVRVCLHVOKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSAFVNETKSTRQTVGLGTQTCETILKQLPNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPHSAFVNETKSTRQTVGLGTQTCETILKQLPNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAFGLRPLVLAEDAQRFTALPPFKKNGCNDNICDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFGLRPLVLAEDAQRFTALPPFKKNGCNDNICDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVDNDEGDSVKTQVTFPPDLDSYRVSTLONORSQSNRLACESASSTEV 840
DB 797 GPREFNVTVDNDEGDSVKTQVTFPPDLDSYRVSTLONORSQSNRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 900
DB 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 916
QY 901 QLELPVKYAVMVTSHGVSSTYKLAFTASENTSRVWQHQQYQVSNLQORSLPISLFLVPV 960
DB 917 QLELPVKYAVMVTSHGVSSTYKLAFTASENTSRVWQHQQYQVSNLQORSLPISLFLVPV 976
QY 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVWNCISIAVCORIQCDIP 1020
DB 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVWNCISIAVCORIQCDIP 1036
QY 1021 PFGIOBEFNATLKNLSFPMYIKTSHNLLIYSTABILFNDVSFTLLPQGGAFVRSQDET 1080
DB 1037 PFGIOBEFNATLKNLSFPMYIKTSHNLLIYSTABILFNDVSFTLLPQGGAFVRSQDET 1096
QY 1081 XVEPFPVNPPLPLIVGSSVGGILLALITALYKLGFFKRYQKDMWSEGGPFGABPQ 1137
DB 1097 XVEPFPVNPPLPLIVGSSVGGILLALITALYKLGFFKRYQKDMWSEGGPFGABPQ 1153

RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5470953ei Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr, Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-286-889-3

Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTEANMTFOENARGFGQSVVQLOGSRVVVGAPOEIVAAQNRGSLYQCDYSTGSCPEI 60
DB 17 FNLDTEANMTFOENARGFGQSVVQLOGSRVVVGAPOEIVAAQNRGSLYQCDYSTGSCPEI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVVKGLCFPLGNSLRQOPQK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVVKGLCFPLGNSLRQOPQK 136
QY 121 PPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTSLFSLMOYSEEF 180
DB 137 PPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTSLFSLMOYSEEF 196
QY 181 RHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFTNGARKNAKILILI 240
DB 197 RHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFTNGARKNAKILIVI 256
QY 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVFQIN 300
DB 257 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVFQIN 316
QY 301 NFEALKTIONLREKIFAIECTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIONLREKIFAIECTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376

361 GVFLYTSKEKSTFNMTRVSDMNDAYLGAAALILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db GVFLYTSKEKSTFNMTRVSDMNDAYLGAAALILNRVQSLVGLGAPRYQHIGLVAMFR 436
Qy QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQVSCPL 480
Db QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQVSCPL 496
Qy PRGQARWQCDVAVLYGEOGQWGRFGAALTVDLVGNGDKLTDVAIGAPGEDNRGAVLYF 540
Db PRGQARWQCDVAVLYGEOGQWGRFGAALTVDLVGNGDKLTDVAIGAPGEDNRGAVLYF 556
Qy HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy PVLRVKATMEFNPREVARVNECDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVT 660
Db PVLRVKATMEFNPREVARVNECDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVT 676
Qy YDLALDSGRPHSRAVNETKSTRQTVGLGTOTCETLKLQLENCEIDPVSPVILRLNF 720
Db YDLALDSGRPHSRAVNETKSTRQTVGLGTOTCETLKLQLENCEIDPVSPVILRLNF 736
Qy SLVGTPLSAFGNLRVLAEDAQRLFTALFPFKNGCNDNICODDLSITFSFMSLDCLVVG 780
Db SLVGTPLSAFGNLRVLAEDAQRLFTALFPFKNGCNDNICODDLSITFSFMSLDCLVVG 796
Qy GPREFNVTTVVNDGEDSVRTQVTFPPFLDLSYRVKSTLQNRQSRVRLACESASSTEV 840
Db GPREFNVTTVVNDGEDSVRTQVTFPPFLDLSYRVKSTLQNRQSRVRLACESASSTEV 856
Qy SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENWPRNKTEF 900
Db SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENWPRNKTEF 916
Qy QLELPVKYAVVMVTVSHGVSTKYLAFTASNTSRVMQHOYQVSNLQSRSLPISLVELVPV 960
Db QLELPVKYAVVMVTVSHGVSTKYLAFTASNTSRVMQHOYQVSNLQSRSLPISLVELVPV 976
Qy RLNQTVIMDRPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPVNVNCSIAVCQRIQCIP 1020
Db RLNQTVIMDRPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPVNVNCSIAVCQRIQCIP 1036
Qy PFCIQEENATLKNLSPDWYIKTSHNHLIIVSTAEILFNDVSVFTLLPQCGAFVRSQSTET 1080
Db PFCIQEENATLKNLSPDWYIKTSHNHLIIVSTAEILFNDVSVFTLLPQCGAFVRSQSTET 1096
Qy KVEPPEVPNPLPIVGVSSVGGLLLALITAAALYKLGFFPKRQYKDMWSEGGPPGABPQ 1137
Db KVEPPEVPNPLPIVGVSSVGGLLLALITAAALYKLGFFPKRQYKDMWSEGGPPGABPQ 1153

RESULT 3
US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,618
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32797
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-618-3
Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PNLDTENAMTQENARFGQSVVOLQGRVVGAPQIIVAAQNGSLYQCDYSTGSCPEI 60
Db 17 PNLDTENAMTQENARFGQSVVOLQGRVVGAPQIIVAAQNGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSAAATSPPLLACGPTVHOTCTSENTYVKGCLFLFGSNLRQPOK 120
Db 77 RLQVPVEAVNMSLGLSAAATSPPLLACGPTVHOTCTSENTYVKGCLFLFGSNLRQPOK 136
Qy 121 FPEALRGCPQSDSDIAFLVDGSGSIIIPHDFFRAKEFISTVMEQKKSKTLPSLMQYSEEP 180
Db 137 FPEALRGCPQSDSDIAFLVDGSGSIIIPHDFFRAKEFISTVMEQKKSKTLPSLMQYSEEP 196
Qy 181 RIHFTFKFQNNPNRSLIKETITOLGRTHTATGIRKVVRELFNITNGARNAFKILLI 240
Db 197 RIHFTFKFQNNPNRSLIKETITOLGRTHTATGIRKVVRELFNITNGARNAFKILLI 256
Qy 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSOBLNTVASKPPRDHVFOIN 300
Db 257 TDGKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSOBLNTVASKPPRDHVFOIN 316
Qy 301 NFEALKTIQNLREKIFAIETOTGSSSSPHEHMSQSGFSAITNSGILLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIETOTGSSSSPHEHMSQSGFSAITNSGILLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFNMTRVSDMNDAYLGAAALILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFNMTRVSDMNDAYLGAAALILNRVQSLVGLGAPRYQHIGLVAMFR 436
Qy 421 QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQVSCPL 480
Db 437 QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQVSCPL 496
Qy 481 PRGQARWQCDVAVLYGEOGQWGRFGAALTVDLVGNGDKLTDVAIGAPGEDNRGAVLYF 540
Db 497 PRGQARWQCDVAVLYGEOGQWGRFGAALTVDLVGNGDKLTDVAIGAPGEDNRGAVLYF 556
Qy 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600

557 HGTSGSGISPSHSQRIAGSKLSPQLQYFGQSLGGQDLTMDGLVDTVGAQGHVLLRSQ 616
601 PVLVRKAIEMFNPREVARNVFECNDQVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660
617 PVLVRKAIEMFNPREVARNVFECNDQVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 676
661 YDLALDSGRPHSAVRNFKNSRTRQTQVLGLTQTCETLKLQLPNCIEDPVPSPVLRNF 720
677 YDLALDSGRPHSAVRNFKNSRTRQTQVLGLTQTCETLKLQLPNCIEDPVPSPVLRNF 736
721 SLVGTPLSAFGNLRPVLAEDAQRLFTALPPEKNGCNDNICODDLSITFSFMSLCLVVG 780
737 SLVGTPLSAFGNLRPVLAEDAQRLFTALPPEKNGCNDNICODDLSITFSFMSLCLVVG 796
781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYKSTVLQNRQSRWSLACESASSTEV 840
797 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYKSTVLQNRQSRWSLACESASSTEV 856
841 SGALKTSKCSINHPIEPENSEVFNITFDVDSKASIGNKLLKANVTSNNNPRNTKTEF 900
857 SGALKTSKCSINHPIEPENSEVFNITFDVDSKASIGNKLLKANVTSNNNPRNTKTEF 916
901 QLELPVKYAVYMYVTSHGVSTKYLNFTASNTSRVMQHQYQVSNLQORSPLISLVLFPV 960
917 QLELPVKYAVYMYVTSHGVSTKYLNFTASNTSRVMQHQYQVSNLQORSPLISLVLFPV 976
961 RLNQTVWRDPQVTFSENLSSTCHTKERLPFSDHFLAELRKAPVNCSTAVCQRIQCDIP 1020
977 RLNQTVWRDPQVTFSENLSSTCHTKERLPFSDHFLAELRKAPVNCSTAVCQRIQCDIP 1036
1021 FFGIQBEFNATLKNLSFQWYIKTSNNHLLIVSTAEILFNDVSFTLLPGQGFVRSQTEF 1080
1037 FFGIQBEFNATLKNLSFQWYIKTSNNHLLIVSTAEILFNDVSFTLLPGQGFVRSQTEF 1096
1081 KVBPFFVNPPLIVGVSSVGLLLALITAAALYKLGFFFRQYKDMMSBGGPPGABEQ 1137
1097 KVBPFFVNPPLIVGVSSVGLLLALITAAALYKLGFFFRQYKDMMSBGGPPGABEQ 1153

RESULT 4

US-08-362-652-3
; Sequence 3, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent-In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.

; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-362-652-3

Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTQENARFGQSVVOLQGSRRVVVGAPQEIVAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTQENARFGQSVVOLQGSRRVVVGAPQEIVAANQSGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTPPQLACGPTVHTQTCSENTYVKGLCFLFSGNLRQOQPK 120
DB 77 RLQVPVEAVNMSLGLSLAATTPPQLACGPTVHTQTCSENTYVKGLCFLFSGNLRQOQPK 136
QY 121 FPEALRGCPQSDSDIAPLVDGSGSIIPHDPRRAKEPISTVMEOLKSKTLFSLMOYSEEF 180
DB 137 FPEALRGCPQSDSDIAPLVDGSGSIIPHDPRRAKEPISTVMEOLKSKTLFSLMOYSEEF 196
QY 181 RIHFTKXEQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
DB 197 RIHFTKXEQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256
QY 241 TDGEKFGPLGVEDVPEADREGVIRYVIGVDAPRSEKSEQLNTVASKPRDRHVFQIN 300
DB 257 TDGEKFGPLGVEDVPEADREGVIRYVIGVDAPRSEKSEQLNTVASKPRDRHVFQIN 316
QY 301 NFEALKTIQNLREKIFAIECTQTGSSSSPHEMSQEGFSAAITNSNGPLLSTVSGSYDWAG 360
DB 317 NFEALKTIQNLREKIFAIECTQTGSSSSPHEMSQEGFSAAITNSNGPLLSTVSGSYDWAG 376
QY 361 GVFLYTSKEKSTFNTWTRVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFNTWTRVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYOHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDLVLICAPHYTPQTRGGQSVSCPL 480
DB 437 QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDLVLICAPHYTPQTRGGQSVSCPL 496
QY 481 PRGQARWQCDAYLYGEGQPGWRFPAALTVLGDVNGDKLTDVAGAEEDNRGAVYLF 540
DB 497 PRGQARWQCDAYLYGEGQPGWRFPAALTVLGDVNGDKLTDVAGAEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPQLQYFGQSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPQLQYFGQSLGGQDLTMDGLVDTVGAQGHVLLRSQ 616
QY 601 PVLVRKAIEMFNPREVARNVFECNDQVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLVRKAIEMFNPREVARNVFECNDQVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSAVRNFKNSRTRQTQVLGLTQTCETLKLQLPNCIEDPVPSPVLRNF 720
DB 677 YDLALDSGRPHSAVRNFKNSRTRQTQVLGLTQTCETLKLQLPNCIEDPVPSPVLRNF 736
QY 721 SLVGTPLSAFGNLRPVLAEDAQRLFTALPPEKNGCNDNICODDLSITFSFMSLCLVVG 780
DB 737 SLVGTPLSAFGNLRPVLAEDAQRLFTALPPEKNGCNDNICODDLSITFSFMSLCLVVG 796
QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYKSTVLQNRQSRWSLACESASSTEV 840

FACTS AND THE ART

RESULTS

US-08-605-672-3
Sequence 3, Application US/08605672
Patent No. 5817515
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5817515e1 Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,672
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-605-672-3

QY 1021 PFGIQEEFNATLKGNSLSPWYIKTNSHLLIIVSTAEILFNDVSFTLLPQCGAFVRSQTEF 1080
Db 1037 PFGIQEEFNATLKGNSLSPWYIKTNSHLLIIVSTAEILFNDVSFTLLPQCGAFVRSQTEF 1096
QY 1081 KVEPPEVNPPLIVGSSVGGILLALITAAALYKLGFFKQYKDMXSEGGPPGABPQ 1137
Db 1097 KVEPPEVNPPLIVGSSVGGILLALITAAALYKLGFFKQYKDMXSEGGPPGABPQ 1153

RESULT 6
US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Wieren, Monica
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-482-293A-3
Query Match 99.8%; Score 5862; DB 2; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLDTENAMTPOENARGFGQSVVQLQGRVVVVGAPQEIIVANQGRSLYQCDYSTGSCBPI 60
Db 17 FNLDTENAMTPOENARGFGQSVVQLQGRVVVVGAPQEIIVANQGRSLYQCDYSTGSCBPI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHCTSCENTYVKGCLFLGSLNLRQQPQK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHCTSCENTYVKGCLFLGSLNLRQQPQK 136
QY 121 PFEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFISTVMEQLKSKTLFSLMOYSEEF 180

Db 137 PFEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFISTVMEQLKSKTLFSLMOYSEEF 196
QY 181 RIHTFKEFQNNPRLSIKPIKLTQGLRTHATGIRKVVRELFINIENGARNAKILILI 240
Db 197 RIHTFKEFQNNPRLSIKPIKLTQGLRTHATGIRKVVRELFINIENGARNAKILIVI 256
QY 241 TDGSKFGDPLGYEDVPEADREBGRVIRVIGVDGAFRSEKSHQELNLTASKPRPHVFOIN 300
Db 257 TDGSKFGDPLGYEDVPEADREBGRVIRVIGVDGAFRSEKSHQELNLTASKPRPHVFOIN 316
QY 301 NFEALKTIONOLREKI PAIEGTQSGSSSPHEMSORGFSAITNGPLLTGVSIDWAG 360
Db 317 NFEALKTIONOLREKI PAIEGTQSGSSSPHEMSORGFSAITNGPLLTGVSIDWAG 376
QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLVGAALAILNRVQSLVLGAPRYQHIGLAVMPR 420
Db 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLVGAALAILNRVQSLVLGAPRYQHIGLAVMPR 436
QY 421 QNTGWESNANVKGTQIGAVFGASLCSVDVDSNSTDLVLIGAPHYVEOTEGGVSVCP 480
Db 437 QNTGWESNANVKGTQIGAVFGASLCSVDVDSNSTDLVLIGAPHYVEOTEGGVSVCP 496
QY 481 PRGORARWQCDVAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNREGAVYLF 540
Db 497 PRGORARWQCDVAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNREGAVYLF 556
QY 541 KTSSTSGISPSHSQRIAGSKSLPRLQYFGQSLSGGQDLTMDGLVDLTVGAQHVLILRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKSLPRLQYFGQSLSGGQDLTMDGLVDLTVGAQHVLILRSQ 616
QY 601 PVLAVKALMBENPREVARNVPECNDQVVKGEAGEVVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLAVKALMBENPREVARNVPECNDQVVKGEAGEVVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAVFETKNSRTRQVGLTQCTETLKLQLENCIEDPVSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAVFETKNSRTRQVGLTQCTETLKLQLENCIEDPVSPVILRLNF 736
QY 721 SLVGTPLSAFQNLRPVLAEDAQRALTALFPPEKNCNDNICODDLSITFFSPMSLDCLVVG 780
Db 737 SLVGTPLSAFQNLRPVLAEDAQRALTALFPPEKNCNDNICODDLSITFFSPMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900
Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 916
QY 901 QLELPVKYAVYVTVVTSRGTSTKYLNFNTASENTSRVMQHQYQVSNLQGRSLPISLVFLVPV 960
Db 917 QLELPVKYAVYVTVVTSRGTSTKYLNFNTASENTSRVMQHQYQVSNLQGRSLPISLVFLVPV 976
QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDPLAELRKAPVNVNCISVACQRIQCDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDPLAELRKAPVNVNCISVACQRIQCDIP 1036
QY 1021 PFGIQEEFNATLKGNSLSPWYIKTNSHLLIIVSTAEILFNDVSFTLLPQCGAFVRSQTEF 1080
Db 1037 PFGIQEEFNATLKGNSLSPWYIKTNSHLLIIVSTAEILFNDVSFTLLPQCGAFVRSQTEF 1096
QY 1081 KVEPPEVNPPLIVGSSVGGILLALITAAALYKLGFFKQYKDMXSEGGPPGABPQ 1137
Db 1097 KVEPPEVNPPLIVGSSVGGILLALITAAALYKLGFFKQYKDMXSEGGPPGABPQ 1153

RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,363
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
TYPE: amino acid
LENGTH: 1153 amino acids
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-363-3

Query Match		99.8%	Score 5862;	DB 2;	Length 1153;
Best Local Similarity		99.3%	Pred. No. 0;		
Matches 1129;		Conservative	7;	Mismatches	1;
				Indels	0;
				Gaps	0;
Qy	1	FNLDENAMTFQENARGFGQSVVQLQGSRRVVVGGAPQBIIVAAQNSGLYQCDYSTGSCPEI	60		
Db	17	FNLDENAMTFQENARGFGQSVVQLQGSRRVVVGGAPQBIIVAAQNSGLYQCDYSTGSCPEI	76		
Qy	61	RLOVPVAVNMSLGLSLAATSPOLLACGPTVHQTSENTRYKGLCFLFGSNLRQOPQK	120		
Db	77	RLOVPVAVNMSLGLSLAATSPOLLACGPTVHQTSENTRYKGLCFLFGSNLRQOPQK	136		
Qy	121	FPBALGCPQSDIAFLVDGSGIIPHDSEKAEFTVMEQLKSKTLFSLMOYSEEF	180		
Db	137	FPBALGCPQSDIAFLVDGSGIIPHDSEKAEFTVMEQLKSKTLFSLMOYSEEF	196		
Qy	181	RIHFTFEQNNPNSRLIKEITQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI	240		
Db	197	RIHFTFEQNNPNSRLIKEITQLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI	256		
Qy	241	TDGKFGDPLGYEDVIEADREGVIRVIVGVDFAFRSEKSRQELNTVASKPRDHVQIN	300		
Db	257	TDGKFGDPLGYEDVIEADREGVIRVIVGVDFAFRSEKSRQELNTVASKPRDHVQIN	316		
Qy	301	NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQSGFAAITSNGPLLSTVGSYDWAQ	360		
Db	317	NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQSGFAAITSNGPLLSTVGSYDWAQ	376		

Qy	361	GVFLYTSKSKSTP	INMTRVDSMDNDAYLGAAAILLNRVQSLVLAGPRYQHIGLVAMER	420
Db	377	GVFLYTSKSKSTP	INMTRVDSMDNDAYLGAAAILLNRVQSLVLAGPRYQHIGLVAMER	436
Qy	421	QNTGMWESNANVRGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYTQETFGGQSVCP	480	
Db	437	QNTGMWESNANVRGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYTQETFGGQSVCP	496	
Qy	481	PRGORARWQCDAYLYGBOGQPNCRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVYLF	540	
Db	497	PRGORARWQCDAYLYGBOGQPNCRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVYLF	556	
Qy	541	HGTSGSGISPSHSOR	TAGSKLSPRLQVFGQSLSGGQDLTMDGLVLDLTGGAQGHVLLRSQ	600
Db	557	HGTSGSGISPSHSOR	TAGSKLSPRLQVFGQSLSGGQDLTMDGLVLDLTGGAQGHVLLRSQ	616
Qy	601	PVLVRKALMEFNPREVARNVFECDQVWVKEAGEVRVCLHWKSTDRDLREGIQSVVT	660	
Db	617	PVLVRKALMEFNPREVARNVFECDQVWVKEAGEVRVCLHWKSTDRDLREGIQSVVT	676	
Qy	661	YDLALDSGRPHSRAVFNENKSTRRTQVGLTQTCTETLKLQLPNCIEDPVSPIVLRNF	720	
Db	677	YDLALDSGRPHSRAVFNENKSTRRTQVGLTQTCTETLKLQLPNCIEDPVSPIVLRNF	736	
Qy	721	SLVGTPLSAPGNLRPVLAEDAQRLFTALPPFKNCGNDNICQDDLSITFSFMSLDCLVVG	780	
Db	737	SLVGTPLSAPGNLRPVLAEDAQRLFTALPPFKNCGNDNICQDDLSITFSFMSLDCLVVG	796	
Qy	781	GPREFNVTVVRNDGSDSYRTQVTFPPDLDSYRKVSTLQNSQSRWRLACESASSTEV	840	
Db	797	GPREFNVTVVRNDGSDSYRTQVTFPPDLDSYRKVSTLQNSQSRWRLACESASSTEV	856	
Qy	841	SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGKLLKLLKANVTSENKMPRTNKTEP	900	
Db	857	SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGKLLKLLKANVTSENKMPRTNKTEP	916	
Qy	901	QLELPVKYAVYVTVTSHGVSTKYLNFTASNTSRVWQHGVQVSNLQORSILPISLVFLVPV	960	
Db	917	QLELPVKYAVYVTVTSHGVSTKYLNFTASNTSRVWQHGVQVSNLQORSILPISLVFLVPV	976	
Qy	961	RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSDFLAELRKAPVNVNCIAVCQRIQCDIP	1020	
Db	977	RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSDFLAELRKAPVNVNCIAVCQRIQCDIP	1036	
Qy	1021	FFGIQEEFNATLKGNSLFDWYIKTSINHLITVSTABILNDSVFTLLPGCGAFVRSQET	1080	
Db	1037	FFGIQEEFNATLKGNSLFDWYIKTSINHLITVSTABILNDSVFTLLPGCGAFVRSQET	1096	
Qy	1081	KVEPFEVNPPLPLTVGSSVGGILLALITAAALYKLGFFKQYKDMSEGPPGABPO	1137	
Db	1097	KVEPFEVNPPLPLTVGSSVGGILLALITAAALYKLGFFKQYKDMSEGPPGABPO	1153	
RESULT 8				
US-09-193-043-3				
Sequence 3, Application US/09193043				
Patent No. 6251395				
GENERAL INFORMATION:				
APPLICANT: Gallatin, Michael W.				
APPLICANT: Van der Vieren, Monica				
TITLE OF INVENTION: No. 6251395el Human 2				
FILE REFERENCE: 27866/35004				
CURRENT APPLICATION NUMBER: US/09/193,043				
CURRENT FILING DATE: 1998-11-16				
EARLIER APPLICATION NUMBER: 08/173,497				
EARLIER FILING DATE: 1993-12-23				
EARLIER APPLICATION NUMBER: 08/286,889				
EARLIER FILING DATE: 1994-08-05				
EARLIER APPLICATION NUMBER: 08/362,652				
EARLIER FILING DATE: 1994-12-21				
EARLIER APPLICATION NUMBER: 08/943,363				
EARLIER FILING DATE: 1997-10-03				


```
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-193-043-3

Query Match      99.8%; Score 5862; DB 3; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQGGSRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQGGSRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSGLSLAATTPPOLACGPTVHQTCSNTYVKGCLFLGSLNLRQPOK 120
DB 77 RLQVPVEAVNMSGLSLAATTPPOLACGPTVHQTCSNTYVKGCLFLGSLNLRQPOK 136

QY 121 FPEALRGCPQSDSDIAFLVDGSGSIIIPHDRAKEFIPTVMQKSKTLPSLMQYSEEF 180
DB 137 FPEALRGCPQSDSDIAFLVDGSGSIIIPHDRAKEFIPTVMQKSKTLPSLMQYSEEF 196

QY 181 RHFTFKFQNNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
DB 197 RHFTFKFQNNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256

QY 241 TDGEKFGDPLGYEDVIPEADREBGIIVYGVGDAPFRSEKSRQELNMTVASKPPDRHVFQIN 300
DB 257 TDGEKFGDPLGYEDVIPEADREBGIIVYGVGDAPFRSEKSRQELNMTVASKPPDRHVFQIN 316

QY 301 NFEALKTIONQUREKIFALEGTQCTSSSFHEMSQEGFSAITNSGPIITVGSYDNAG 360
DB 317 NFEALKTIONQUREKIFALEGTQCTSSSFHEMSQEGFSAITNSGPIITVGSYDNAG 376

QY 361 GYPLVTSKSTFNNTRVDSMDNAYLGAYAAAILRNVRQSLVLGAPRYQHIGLVAMPR 420
DB 377 GYPLVTSKSTFNNTRVDSMDNAYLGAYAAAILRNVRQSLVLGAPRYQHIGLVAMPR 436

QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLICAPHYVQTRGQVSVCP 480
DB 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLICAPHYVQTRGQVSVCP 496

QY 481 PRGQARWQCDVLYGEGQPPGRFGAALTVLGVNGDKLTDVAIGAPCEEDNRGAVYLF 540
DB 497 PRGQARWQCDVLYGEGQPPGRFGAALTVLGVNGDKLTDVAIGAPCEEDNRGAVYLF 556

QY 541 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLWGAQGHVLIIRSQ 600
DB 557 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLWGAQGHVLIIRSQ 616

QY 601 PVLRYKAIEMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLRYKAIEMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676

QY 661 YDLALDSGRPHSAFNETXSTRQTQVLGHTQTCETLKLQIPNCIEDPUSPIVLRNLF 720
DB 677 YDLALDSGRPHSAFNETXSTRQTQVLGHTQTCETLKLQIPNCIEDPUSPIVLRNLF 736

QY 721 SLVGTPLSAFGNLRPVLAEDAORLFTALPPFEKNGCNDNI CODDLSTIFSMSLDCLVVG 780
DB 737 SLVGTPLSAFGNLRPVLAEDAORLFTALPPFEKNGCNDNI CODDLSTIFSMSLDCLVVG 796

QY 781 GPREFNVTVTVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNSQSRQSWRLACESASSTEV 840
DB 797 GPREFNVTVTVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNSQSRQSWRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIIPENSEVTNITFDVDSKASLGNKLLKANVTSENNPRNKTEF 900
DB 857 SGALKSTSCSINHPIIPENSEVTNITFDVDSKASLGNKLLKANVTSENNPRNKTEF 916

QY 901 QLELPVKYAVYVWVTVSHGVSTKYLNTASENTSRVMOHQYQVSNLQGRSLPISLVFLVPV 960
```

Db 317 NFEALKTIONQREKIPAIETGQTGSSSSFEHMSQEGFSAAITSGPILLSVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFIMNTVSDMDNDAYLGAAAIILNRVQSLVILGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFIMNTVSDMDNDAYLGAAAIILNRVQSLVILGAPRYOHIGLVAMFR 436
Qy 421 QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDVILGAPHYEQTREGQVSVCP 480
Db 437 QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDVILGAPHYEQTREGQVSVCP 496
Qy 481 PRGQARWQCDVLYGEGQPGWRFGAALTIVLDVNGDKLTVDVAGPGEEDNRGAVYLP 540
Db 497 PRGQARWQCDVLYGEGQPGWRFGAALTIVLDVNGDKLTVDVAGPGEEDNRGAVYLP 556
Qy 541 HGTSGSGISPSHSORIASGKLSPLQYFGQSLSGQDLTMDGLVDLTGACQHVLLRSQ 600
Db 557 HGTSGSGISPSHSORIASGKLSPLQYFGQSLSGQDLTMDGLVDLTGACQHVLLRSQ 616
Qy 601 PVLRVKAIMENPREVARNVPECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIMENPREVARNVPECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQLPNCIEDPSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQLPNCIEDPSPVILRLNF 736
Qy 721 SLVGTPLSAFQNLRLPVLAEADAQRLFTALFPFEKNCNDNICQDDLSITFSPMSLDCLVVG 780
Db 737 SLVGTPLSAFQNLRLPVLAEADAQRLFTALFPFEKNCNDNICQDDLSITFSPMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLISYKXSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLISYKXSTLQNRQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTENITFQVDSKASLGNKLLKANVTSENMPRTNKTET 900
Db 900 SGALKSTSCSINHPIPPENSEVTENITFQVDSKASLGNKLLKANVTSENMPRTNKTET 916
Qy 901 QLELPVKAVVMVTVSHGVSTYKLYNFTASENTSRVMQHQYQVSNLQGRSLPISLFLPV 960
Db 917 QLELPVKAVVMVTVSHGVSTYKLYNFTASENTSRVMQHQYQVSNLQGRSLPISLFLPV 976
Qy 961 RINQTVIWRDPVOTVSENLSSCTCHKERLPKSHDFLAEIRKAPVVCISIAVCQRIQCDIP 1020
Db 977 RINQTVIWRDPVOTVSENLSSCTCHKERLPKSHDFLAEIRKAPVVCISIAVCQRIQCDIP 1036
Qy 1021 PFGIOEEFNATLKGNSLSDWYIKTSHNLLIVSTAEILLFNDVSFTLLPQCGAFVRSQDET 1080
Db 1037 PFGIOEEFNATLKGNSLSDWYIKTSHNLLIVSTAEILLFNDVSFTLLPQCGAFVRSQDET 1096
Qy 1081 KVEPEVENPPLIVGSSVGGILLALITAAALYKLGFFRKQYKQNMSSGGPPGABEQ 1137
Db 1097 KVEPEVENPPLIVGSSVGGILLALITAAALYKLGFFRKQYKQNMSSGGPPGABEQ 1153
RESULT 10
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. 6620915
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6620915e1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652

; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.8%; Score 5862; DB 4; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ENLDTENAMTFQENARGFGQSVVQLOQSRVVVGAPOIIVAANQORGSILYQCDYSTGSCPEI 60
Db 17 ENLDTENAMTFQENARGFGQSVVQLOQSRVVVGAPOIIVAANQORGSILYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCESTYVKGCLFPGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCESTYVKGCLFPGSNLRQOPQK 136
Qy 121 FPEALRCPOBDSIAELVDGSGSIIIPHDERRAKEPISTVMEQLKXSKTLPFLMOYSEEP 180
Db 137 FPEALRCPOBDSIAELVDGSGSIIIPHDERRAKEPISTVMEQLKXSKTLPFLMOYSEEP 196
Qy 181 RIHFTPEFQNNPNRSLIRKPIITOLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 240
Db 197 RIHFTPEFQNNPNRSLIRKPIITOLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256
Qy 241 TDGEKFDPLGYEDVIEADREGVIRVYVGVGDAPRSEKROELNTVASKPPRDHVQIN 300
Db 257 TDGEKFDPLGYEDVIEADREGVIRVYVGVGDAPRSEKROELNTVASKPPRDHVQIN 316
Qy 301 NFEALKTIQNLREKIPAIETGQTGSSSSFEHMSQEGFSAAITSGPILLSVGSYDWAG 360
Db 317 NFEALKTIQNLREKIPAIETGQTGSSSSFEHMSQEGFSAAITSGPILLSVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFIMNTVSDMDNDAYLGAAAIILNRVQSLVILGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFIMNTVSDMDNDAYLGAAAIILNRVQSLVILGAPRYOHIGLVAMFR 436
Qy 421 QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDVILGAPHYEQTREGQVSVCP 480
Db 437 QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDVILGAPHYEQTREGQVSVCP 496
Qy 481 PRGQARWQCDVLYGEGQPGWRFGAALTIVLDVNGDKLTVDVAGPGEEDNRGAVYLP 540
Db 497 PRGQARWQCDVLYGEGQPGWRFGAALTIVLDVNGDKLTVDVAGPGEEDNRGAVYLP 556
Qy 541 HGTSGSGISPSHSORIASGKLSPLQYFGQSLSGQDLTMDGLVDLTGACQHVLLRSQ 600
Db 557 HGTSGSGISPSHSORIASGKLSPLQYFGQSLSGQDLTMDGLVDLTGACQHVLLRSQ 616
Qy 601 PVLRVKAIMENPREVARNVPECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIMENPREVARNVPECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQLPNCIEDPSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQLPNCIEDPSPVILRLNF 736
Qy 721 SLVGTPLSAFQNLRLPVLAEADAQRLFTALFPFEKNCNDNICQDDLSITFSPMSLDCLVVG 780
Db 737 SLVGTPLSAFQNLRLPVLAEADAQRLFTALFPFEKNCNDNICQDDLSITFSPMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLISYKXSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLISYKXSTLQNRQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTENITFQVDSKASLGNKLLKANVTSENMPRTNKTET 900
Db 900 SGALKSTSCSINHPIPPENSEVTENITFQVDSKASLGNKLLKANVTSENMPRTNKTET 916

Db 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEP 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNTASENTSRVMQHQYQVSNLQSRSLPISLVFLV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNTASENTSRVMQHQYQVSNLQSRSLPISLVFLV 976
Qy 961 RLNOVTWDRPOVTFSENLSTCHTKERLPKSHDPLAELRKAPVWNCISAVCORIQCDIP 1020
Db 977 RLNOVTWDRPOVTFSENLSTCHTKERLPKSHDPLAELRKAPVWNCISAVCORIQCDIP 1036
Qy 1021 PFGIOBEFNATLKGMLSPDWYIKTSHNHLIIVSTAEILLFNDVSFTLLPQCGAFVRSQTET 1080
Db 1037 PFGIOBEFNATLKGMLSPDWYIKTSHNHLIIVSTAEILLFNDVSFTLLPQCGAFVRSQTET 1096
Qy 1081 KVEPFEVNPPLIVGSSVGGLLALITAALYKLGFFKQYKQVMSGGPPGAEPQ 1137
Db 1097 KVEPFEVNPPLIVGSSVGGLLALITAALYKLGFFKQYKQVMSGGPPGAEPQ 1153

RESULT 11
US-08-476-062A-43
; Sequence 43, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212,573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-476-062A-43

Query Match 99.2%; Score 5831.5; DB 2; Length 1152;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFOENARGFGOSVVOLOQSRVVGAPQBIIVAAQORGSLYOCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGOSVVOLOQSRVVGAPQBIIVAAQORGSLYOCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATSPQLACGPTVHQTCSNTYVKGCLFPLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATSPQLACGPTVHQTCSNTYVKGCLFPLFGSNLRQOPQK 136
Qy 121 FPEALRCPOBDSIDIAFLVDSGSIIPHDPRRAKEPFTSTWBOLEKSKTILFSLMQYSSEF 180
Db 137 FPEALRCPOBDSIDIAFLVDSGSIIPHDPRRAKEPFTSTWBOLEKSKTILFSLMQYSSEF 196
Qy 181 RIHFTFKFQNNPNRSLIKPIITOLLGRTHATGIRKVVRELFINITNGARNAFKILILI 240
Db 197 RIHFTFKFQNNPNRSLIKPIITOLLGRTHATGIRKVVRELFINITNGARNAFKILILI 256
Qy 241 TDGEKFGDPLGYEDVIPEADREBQVIRVIVGVDAFRSEKSRQELNLTIVASKPPRPHVQIN 300
Db 257 TDGEKFGDPLGYEDVIPEADREBQVIRVIVGVDAFRSEKSRQELNLTIVASKPPRPHVQIN 316
Qy 301 NFEALKTIQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAITNSNGPLISTVGSYDWAG 360
Db 317 NFEALKTIQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAITNSNGPLISTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDADAYLGYAAAILLRNVOSLVLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDADAYLGYAAAILLRNVOSLVLGAPRYOHIGLVAMFR 436
Qy 421 QNTGWTESNANVKCTQIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYEQTGGQVSVCP 480
Db 437 QNTGWTESNANVKCTQIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYEQTGGQVSVCP 496
Qy 481 PRGORARWQCDVLYGBQGPWGFAGALTIVLGDVNGDKLTDVAIGAPGEDNAGAVL 540
Db 497 PRG-RARWQCDVLYGBQGPWGFAGALTIVLGDVNGDKLTDVAIGAPGEDNAGAVL 555
Qy 541 HGTSGSISPSHSORISAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 556 HGTSGSISPSHSORISAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 615
Qy 601 PVLRVKALMEFNPREVARNFECDQVWVKGAEGRVCLHVQKSTRDLREGQIQSVVT 660
Db 616 PVLRVKALMEFNPREVARNFECDQVWVKGAEGRVCLHVQKSTRDLREGQIQSVVT 675
Qy 661 YDLALDSGRPHSRVFNETKSTRQOTVLGTCTETLKLQLPNCIEDPVSPIVLRNF 720
Db 676 YDLALDSGRPHSRVFNETKSTRQOTVLGTCTETLKLQLPNCIEDPVSPIVLRNF 735
Qy 721 SLVGTPLSAFGNLRPVLAEADAQRLFTALPPEKNGNDNICODDLSITFSFMSLDCLV 780
Db 736 SLVGTPLSAFGNLRPVLAEADAQRLFTALPPEKNGNDNICODDLSITFSFMSLDCLV 795
Qy 781 GPRESNVTVVNDGEDSVTQVTFEFLDLSYRKVSTLQNRORSORSLACESASSTEV 840
Db 796 GPRESNVTVVNDGEDSVTQVTFEFLDLSYRKVSTLQNRORSORSLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEP 900
Db 856 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEP 915
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNTASENTSRVMQHQYQVSNLQSRSLPISLVFLV 960
Db 916 QLELPVKYAVYVMTSHGVSTKYLNTASENTSRVMQHQYQVSNLQSRSLPISLVFLV 975
Qy 961 RLNOVTWDRPOVTFSENLSTCHTKERLPKSHDPLAELRKAPVWNCISAVCORIQCDIP 1020
Db 976 RLNOVTWDRPOVTFSENLSTCHTKERLPKSHDPLAELRKAPVWNCISAVCORIQCDIP 1035
Qy 1021 PFGIOBEFNATLKGMLSPDWYIKTSHNHLIIVSTAEILLFNDVSFTLLPQCGAFVRSQTET 1080
Db 1036 PFGIOBEFNATLKGMLSPDWYIKTSHNHLIIVSTAEILLFNDVSFTLLPQCGAFVRSQTET 1095
Qy 1081 KVEPFEVNPPLIVGSSVGGLLALITAALYKLGFFKQYKQVMSGGPPGAEPQ 1137

Best Local Similarity 99.0%; Pred. No. 0;
Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGGSRVVVGAPQEIIVAAHQSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGGSRVVVGAPQEIIVAAHQSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSGLSLAATSPOLLACGPTVHTCSENTYVKGCLFPLGSLNRQQPOK 120
DB 77 RLQVPVEAVNMSGLSLAATSPOLLACGPTVHTCSENTYVKGCLFPLGSLNRQQPOK 136

QY 121 FPEARLGCPOBDSIAFLVDGSGSIIPHDPRRAKEFISTVMEOLKSKTLFSLMOYSEEP 180
DB 137 FPEARLGCPOBDSIAFLVDGSGSIIPHDPRRAKEFISTVMEOLKSKTLFSLMOYSEEP 196

QY 181 RIHFTFKFQNNPNSLIPKPTQLLGRTHATGIRKVVRELFTNGARKNAFKILILI 240
DB 197 RIHFTFKFQNNPNSLIPKPTQLLGRTHATGIRKVVRELFTNGARKNAFKILIVI 256

QY 241 TDGKFGDPLGVEDVPEADREGVIRYVIGVDDAFRSEKSRQELMTVASKPRDHVFQIN 300
DB 257 TDGKFGDPLGVEDVPEADREGVIRYVIGVDDAFRSEKSRQELMTVASKPRDHVFQIN 316

QY 301 NFEALKTIONOLREKIFATEGTQSSSFHEHMSQEGFSAATNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIONOLREKIFATEGTQSSSFHEHMSQEGFSAATNSGPLLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSPINNTRVDSMDNAYLGYAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKEKSPINNTRVDSMDNAYLGYAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436

QY 421 QNTGWESNANVKGTOIGAYFCASLCSVDVDSNGSTDVLICAPHYVQTRGGQSVQCLP 480
DB 437 QNTGWESNANVKGTOIGAYFCASLCSVDVDSNGSTDVLICAPHYVQTRGGQSVQCLP 496

QY 481 PRGORARWQDAVLGYEQOQPMRFGAALTVLGVDNGDKLTDVAI GAPGEEDNRGAVYLF 540
DB 497 PRG-RARWQDAVLGYEQOQPMRFGAALTVLGVDNGDKLTDVAI GAPGEEDNRGAVYLF 555

QY 541 HGTSGSISPSHSQRTAGSKLSPRLQYFGOSLISGQDLMGDLVLTGVAQGHVLLRSQ 600
DB 556 HGTSGSISPSHSQRTAGSKLSPRLQYFGOSLISGQDLMGDLVLTGVAQGHVLLRSQ 615

QY 601 PVLRVKAIKMEFNPREVARNVFCNDQVVKGEAGEVRVCLRHVQKSTRDLRSGQIQSVVT 660
DB 616 PVLRVKAIKMEFNPREVARNVFCNDQVVKGEAGEVRVCLRHVQKSTRDLRSGQIQSVVT 675

QY 661 YDLALDSGRPHSRVNETKNTSTRTQVLGTQTCETLKLQLPNCIEDPVSPIVLRNF 720
DB 676 YDLALDSGRPHSRVNETKNTSTRTQVLGTQTCETLKLQLPNCIEDPVSPIVLRNF 735

QY 721 SLVGTPLSAFNLRPVLAEDAQLFTALFFPKNCNDNI CODDLSITPSFMSLDCLVVG 780
DB 736 SLVGTPLSAFNLRPVLAEDAQLFTALFFPKNCNDNI CODDLSITPSFMSLDCLVVG 795

QY 781 GPREFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWPLACESASSTEV 840
DB 796 GPRESNVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWPLACESASSTEV 855

QY 841 SGALKTSCTSIINHPIDPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEP 900
DB 856 SGALKTSCTSIINHPIDPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEP 915

QY 901 QLELPVKYAVVMVTSKYLNFASNTSRVMQHQYQVSNLQORSPLISLFLVFPV 960
DB 916 QLELPVKYAVVMVTSKYLNFASNTSRVMQHQYQVSNLQORSPLISLFLVFPV 975

QY 961 RLNQTWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORQCDDIP 1020
DB 976 RLNQTWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORQCDDIP 1035

QY 1021 FFGIQEENATLKGNSLFDWYIKTSHNHLIVSTABILFNDVSFTLLPQCGAFVRSQTF 1080

DB 1036 FFGIQEENATLKGNSLFDWYIKTSHNHLIVSTABILFNDVSFTLLPQCGAFVRSQTF 1095

QY 1081 KVPFPEVFNPLPIVGVSSVGLLILALITAAALYKLGFFKQYKQMMSEGGPPGAEPO 1137
DB 1096 KVPFPEVFNPLPIVGVSSVGLLILALITAAALYKLGFFKQYKQMMSEGGPPGAEPO 1152

RESULT 14
US-08-476-062A-44
; Sequence 44, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Atrahout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212,573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1163 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-476-062A-44

Query Match 59.0%; Score 3469; DB 2; Length 1163;
Best Local Similarity 61.0%; Pred. No. 5.9e-284;
Matches 689; Conservative 141; Mismatches 293; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGGSRVVVGAPQEIIVAAHQSLYQCDYSTGSCPEI 60
DB 20 FNLDTENAMTFOENARGFGQSVVQLQGGSRVVVGAPQEIIVAAHQSLYQCDYSTGSCPEI 79

QY 61 RLQVPVEAVNMSGLSLAATSPOLLACGPTVHTCSENTYVKGCLFPLGSLNRQQPOK 120
DB 80 RLQVPVEAVNMSGLSLAATSPOLLACGPTVHTCSENTYVKGCLFPLGSLNRQQPOK 137

QY 121 FPEARLGCPOBDSIAFLVDGSGSIIPHDPRRAKEFISTVMEOLKSKTLFSLMOYSEEP 180
DB 138 FPEARLGCPOBDSIAFLVDGSGSIIPHDPRRAKEFISTVMEOLKSKTLFSLMOYSEEP 197

QY 181 RIHFTFKFQNNPNSLIPKPTQLLGRTHATGIRKVVRELFTNGARKNAFKILILI 240

198 QTHFTFEFRNTSPLSLASVHQLQGGFTYTATATQNVVHLFASVYCARDAKILIVI 257
241 TDGEKFGDPLGVEDVPEADREGVIRVYGVGDAPRSEKSEKQELNTVASKPRDHVQIN 300
258 TDGKEGDSLDYKDVIPWADAGIIRYAGVGLAPQNRNSWKLNDIASKPSQEHIFKVE 317
301 NFEALKTQNLREKIPAEIGTQTSSTSSSFEHMSQEGFSAITSNGPILLSVGSYDWAG 360
318 DFDALKDQNLQKKEKIPAEIGTETSSSSFELEMAQEGFSAVTPDGPVLGAVGFTWSG 377
361 GVFLYTSKEKSTFIMNTRVDSMDAYLGAAAIILNRVQSLVGLGAPRYOHTGKAVFT 420
378 GAFLYPPNMSPTFIMNSQENVDMRDSYLGSTELALWGVQSLVGLGAPRYOHTGKAVFT 437
421 QNTGWNESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTGQGVSVCP 480
438 QVSRQWRMAEVTGTQIGSYFGASLCSVDVDTGSTDVLVIGAPHYEQTGQGVSVCP 497
481 PRGQARWCCDVLVYGEQGWGFGAALTVLGVNGDKLTDVAIGAPGBEDNRGAVYLF 540
498 PRGMR-RWMCDAVLYGEQGWGFGAALTVLGVNGDKLTDVAVIGAPGBEDNRGAVYLF 556
541 HTSGSGISPSHSORIASKSLPRLOFQSGISGGDLTMDGLVDTVGAQCHVLLRSQ 600
557 HGVLPSPISPSHSORIASQLSRLOFQSGISGGDLTMDGLVDTVGAQCHVLLRSQ 616
601 PVLRYKAIEMFNPREVARNVFCNDQVVKGEAGVRVCLHVQKSTRDLREGIQSVVT 660
617 PVLWGVSMQFTPAEIPASAFECREQVWSEQTLVQSNICLYDKSKNLLGSRDIQSSVT 676
661 YDLALDSGPHSRAVFNENKSTRQTOVLGTQCTETLKLQPNCEIDPVSPIVLRNF 720
677 LDALDPGLSPRATFQTKNSLSRVRVIGUKAHCEPNLLSPSCVDSVTPITLRNF 736
721 SLVGTPLGAFGNLRVPLAEDARLFTALFPPEKNGCNDNICODULSITFSNLSCLVVG 780
737 TLVGPPLAFNLRPMUAAALQRYFTASLPPEKNGADHICODNLGISFSPGLKSLVG 796
781 GPRBNVTVVNDGEDSVRTQVTFEPLDLSYKVSITLQNSORSNRKLACESASTEV 840
797 SNLENAEVMWVNDGEDSVGTITFESHAPGLSYRYVAEQKQGRSLHLTCDSPVGV-- 854
841 SGALKSTCSINHEPFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEP 900
855 SGTWSTSCINELFRGAQITFLATFDVSPKAVLGBRLITANVSSENTPRTSKTF 914
901 QLELPVKYAVVYVTSYGVSTKYVFTAS-ENTSRVMOHQYQVSNLQSRSLPISLVFLVP 959
915 QLELPVKYAVVYVTSYGVSTKYVFTAS-ENTSRVMOHQYQVSNLQSRSLPISLVFLVP 974
960 VRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKPVNVCSIAVCCRIQCDI 1019
975 VELNQAVMMDVEVSHQPNPSLRCSSEKIAPPASDFLAHQNPVLDSCSIACCLPRCDV 1034
1020 PFFGIOEEFNATLKNLSFDWIKTSHNLLIIVSTAEILPNDSPVFTLLPGQAPVRSQTE 1079
1035 PFSVQGEELDTLKNLSFGWRQILQKVVSVVAEITFDTSVYSQLPQGEAFNRQTT 1094
1080 TKVPEPEVNPPLIVGSSVGGHLLIILALITAYLKLGEKRYQKDMSE 1128
1095 TVLEKYNVHNPTPLIVGSSIGGLLLIILALITAVLYKVGPFKRYKEMEE 1143

RESULT 15
PCT-US96-01314-44
; Sequence 44, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 1163
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US96-01314-44

Query March 59.08; Score 3469; DB 5; Length 1163;
Best Local Similarity 61.04; Pred. No. 5.9e-284; Indels 6; Gaps 4;
Matches 689; Conservative 141; Mismatches 293;

QY 1 FNLDTENAMTFQENARGFGOSVQLOQSVVVGAPQIRIVANQSGSLYQCYSTGSCBPI 60
DB 20 FNLDTBELTAFRVDSAGFGDSVQVANSVVGAPQIRIVANQSGSLYQCYSTGSCBPI 79
QY 61 RLQVPEAVNMSLGLSLAATSPQLACGPTVHOTSENVYVGLCLPFLGSLNRQQPK 120
DB 80 GLQVPEAVNMSLGLSLAATSPQLACGPTVHOTSENVYVGLCLPFLGSLNRQQPK 137
QY 121 FPEALRGCPQSDSDIAFLVDGSGSIIIPHDPRRAKEFISTVMEQLKSKTLFSLMOYSBF 180
DB 138 LPVSRQECPRQEQIIVFLDGSISSRNFAVAVISQFQRPSTQFSLMOYSBF 197
QY 181 RIHTTFEPQNNPNSRLIKITQILGTHATGIRKVVRELFTNAGAKNAKILILI 240
DB 198 QTHFTFEFRNTSPLSLASVHQLQGGFTYTATATQNVVHLFASVYCARDAKILIVI 257
QY 241 TDGEKFGDPLGVEDVPEADREGVIRVYGVGDAPRSEKSEKQELNTVASKPRDHVQIN 300
DB 258 TDGKEGDSLDYKDVIPWADAGIIRYAGVGLAPQNRNSWKLNDIASKPSQEHIFKVE 317
QY 301 NFEALKTQNLREKIPAEIGTQTSSTSSSFEHMSQEGFSAITSNGPILLSVGSYDWAG 360
DB 318 DFDALKDQNLQKKEKIPAEIGTETSSSSFELEMAQEGFSAVTPDGPVLGAVGFTWSG 377
QY 361 GVFLYTSKEKSTFIMNTRVDSMDAYLGAAAIILNRVQSLVGLGAPRYOHTGKAVFT 420
DB 378 GAFLYPPNMSPTFIMNSQENVDMRDSYLGSTELALWGVQSLVGLGAPRYOHTGKAVFT 437
QY 421 QNTGWNESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTGQGVSVCP 480
DB 438 QVSRQWRMAEVTGTQIGSYFGASLCSVDVDTGSTDVLVIGAPHYEQTGQGVSVCP 497
QY 481 PRGQARWCCDVLVYGEQGWGFGAALTVLGVNGDKLTDVAIGAPGBEDNRGAVYLF 540
DB 498 PRGMR-RWMCDAVLYGEQGWGFGAALTVLGVNGDKLTDVAVIGAPGBEDNRGAVYLF 556

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:05:16 ; Search time 14.559 Seconds
(without alignments)
7512.163 Million cell updates/sec

Title: US-09-902-481B-6
Perfect score: 5875
Sequence: 1 FNLTENAMTFQENARFGQ.....PKROYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:
1: PIR1:
2: PIR2:
3: PIR3:
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5868	99.9	1153	1 RWU1B	cell surface glyco
2	4476	76.2	1153	2 S00551	leukocyte surface
3	3483	59.3	1163	1 RWU1C	cell surface glyco
4	1548.5	26.4	1170	2 S03308	cell surface glyco
5	1533.5	26.1	1163	2 I56126	lymphocyte fuction
6	1149	19.6	1179	2 A53213	integrin alpha-E c
7	1102.5	18.8	1151	2 A45226	integrin alpha-1 c
8	1084	18.5	1170	2 I45914	integrin alpha 2 s
9	1072	18.2	1178	2 S44142	VLA-2 protein homo
10	1069	18.2	1181	2 A33998	integrin alpha-2 c
11	1060	18.0	1180	2 A35854	integrin alpha-1 c
12	683	11.3	1039	2 A41131	lymphocyte-Peyer's
13	633	10.8	1038	2 S06046	integrin alpha-4 c
14	630	10.7	1035	2 I58409	integrin alpha-9 c
15	614.5	10.5	1041	2 T31437	integrin alpha cha
16	579.5	9.9	1054	2 JC7294	alhap integrin -
17	572.5	9.7	1051	2 A35761	cell surface glyco
18	567.5	9.7	1053	2 I55534	VLA-3 alpha subuni
19	555.5	9.5	1053	2 S44250	integrin alpha-5 c
20	543.5	9.3	1034	2 A36108	integrin alpha-V c
21	535	9.1	1044	2 T10050	integrin alpha-v c
22	532	9.1	1049	2 A27079	fibronectin recept
23	532	9.1	1073	2 B36429	integrin alpha-6 c
24	530.5	9.0	1072	2 A38457	integrin alpha-6 c
25	529.5	9.0	1051	2 A40021	integrin alpha-3 c
26	526	9.0	1048	2 A27421	integrin alpha-5 c
27	525.5	8.9	1031	2 A41543	integrin alpha-6 c
28	517	8.8	1044	2 S16516	integrin alpha-8 c
29	496.5	8.5	1146	2 S40311	integrin - fruit f

30 495.5 8.4 1039 2 A34269 integrin alpha-2b
31 494 8.4 1394 2 A29637 position-specific
32 490 8.3 1137 2 JC5950 integrin alpha-7 c
33 488 8.3 1135 2 I61186 alpha-7 integrin -
34 486 8.3 126 2 B30892 leukocyte adhesion
35 485.5 8.3 1037 2 S38783 glycoprotein IIb -
36 469.5 8.0 1106 2 A60163 integrin alpha cha
37 466 7.9 1226 2 S44924 RS4F2.1 protein -
38 454 7.7 1045 2 S60571 integrin alpha v c
39 445.5 7.6 1139 2 S28277 hypothetical prote
40 422 7.2 1115 2 T09403 integrin alpha cha
41 417.5 7.1 1115 2 T09433 integrin alpha cha
42 391 6.7 764 2 I36316 glycoprotein IIb -
43 309 5.3 1086 2 T18523 integrin alpha cha
44 299 5.1 604 2 I36917 glycoprotein IIb -
45 296.5 5.0 272 2 A55348 integrin alpha-1 -

ALIGNMENTS

RESULT 1

RWU1B
cell surface glycoprotein CD11b precursor [validated] - human
N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Ma
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000
C:Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; 152567
R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.
J. Biol. Chem. 263, 12403-12411, 1988
A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CR
B.
A:Reference number: A31108; MUID:88315033; PMID:2457584
A:Accession: A31108
A:Molecule type: mRNA
A:Residues: 1-1153 <COR>
A:CROSS-references: GB:J03925; NID:g187284; PIDN:AAA59544.1; PID:g307148
A>Note: Part of this sequence was confirmed by protein sequencing
R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.
J. Cell Biol. 106, 2153-2158, 1988
A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor b
A:Reference number: A28915; MUID:88257215; PMID:2454931
A:Accession: A28915
A:Molecule type: mRNA
A:Residues: 1-499,501-965, 'P', 967-1153 <ARN>
A:CROSS-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:g186935; PIDN:AAA594
A>Note: the authors translated the codon TAC for residue 1129 as Thr
A>Note: part of this sequence, including the amino end of the mature protein, was confir
R:Shelley, C.S.; Arnaout, M.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991
A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally reg
A:Reference number: A41600; MUID:92073318; PMID:1683702
A:Accession: A41600
A:Molecule type: DNA
A:Residues: 1-9 <SHE>
A:CROSS-references: GB:M76724; NID:g180018; PIDN:AAA58410.1; PID:g553215
R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988
A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesi
A:Reference number: A34193; MUID:88190151; PMID:2833753
A:Accession: A30892
A:Molecule type: mRNA
A:Residues: 917-1042 <AR2>
A:CROSS-references: GB:M18044
R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989
A:Title: CDNA sequence for the alphaM subunit of the human neutrophil adherence receptor
A:Reference number: A32218; MUID:89098893; PMID:2563162
A:Accession: A32218
A:Molecule type: mRNA
A:Residues: 9-1153 <HIC>
A:CROSS-references: GB:J04145; NID:g189068; PIDN:AAA59903.1; PID:g386975

A>Note: part of this sequence was confirmed by protein sequencing
 J. Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
 J. Immunol. 150, 480-490, 1993
 A>Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
 n during evolution.
 A:Reference number: A46526; MUID:93123748; PMID:8419480
 A:Accession: A46526
 A>Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-499,501-1153 <FLE>
 A:Cross-references: GB:552227; NID:9263047; PIDN:AA24821.1; PID:9263049
 A>Note: the last three bases of intron 13, CAG, are included in some but not all mature
 A>Note: sequence extracted from NCBI backbone (NCBIP:121963)
 R.Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
 Biochim. Biophys. Acta 874, 368-371, 1986
 A>Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
 A:Reference number: A30664; MUID:87076671; PMID:3539202
 A:Accession: A26091
 A:Molecule type: protein
 A:Residues: 17-31 <PIE>
 A:Experimental source: granulocytes
 R.Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
 Blood 79, 865-870, 1992
 A>Title: Characterization of the myeloid-specific CD11b promoter.
 A:Reference number: I52567; MUID:92144986; PMID:1346576
 A:Accession: I52567
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-9 <RES>
 A:Cross-references: GB:M84477; NID:9180184; PIDN:AA51960.1; PID:9553219
 C:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
 C:Genetics:
 A:Gene: ITGAM; CR3A
 A:Cross-references: GDB:120599; OMIM:120980
 A:Map position: 16p11.2-16p11.2
 A>Note: promoter contains a GATA motif and two Spl consensus binding sites
 C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homod
 C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
 F:17-1108/Domain: extracellular #status predicted <EXT>
 F:148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
 F:465-473/Region: calcium/magnesium binding #status predicted
 F:530-538/Region: calcium/magnesium binding #status predicted
 F:593-601/Region: calcium/magnesium binding #status predicted
 F:1109-1134/Domain: transmembrane #status predicted <TM>
 F:1135-1153/Domain: intracellular #status predicted <INT>
 F:86,240,391,469,693,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding
 Query Match 99.9%; Score 5868; DB 1; Length 1153;
 Best Local Similarity 99.6%; Pred. NO. 0;
 Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ENLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAVNAQSGSLYQCDYSTGSCPEI 60
 DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAVNAQSGSLYQCDYSTGSCPEI 76
 QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQCSENTYVKGCLFLFGSLNLRQPOK 120
 DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQCSENTYVKGCLFLFGSLNLRQPOK 136
 QY 121 FPEALRGCPQEDSDIAFLDGGSSIIIPHPFRMKPEVSTVMEQLKSKTLFSLMOYSEEF 180
 DB 137 FPEALRGCPQEDSDIAFLDGGSSIIIPHPFRMKPEVSTVMEQLKSKTLFSLMOYSEEF 196
 QY 181 RIHFTKFEQNNPNRSLVKPIITOLLGRTHATGVAKVIRELLINTNGARKNAKILVI 240
 DB 197 RIHFTKFEQNNPNRSLVKPIITOLLGRTHATGVAKVIRELLINTNGARKNAKILVI 256
 QY 241 TDGEKFGDPLGYEDVTPADREGVIRVIGVGAPFSEKSRQELNTIASKPRDHVQVN 300
 DB 257 TDGEKFGDPLGYEDVTPADREGVIRVIGVGAPFSEKSRQELNTIASKPRDHVQVN 316

QY 301 NFPAALXTIONQIREKI FAIEGTCTGSSSSPEHEMSQSGFSAATTSNGPLLSTVGSYDWAG 360
 DB 317 NFPAALXTIONQIREKI FAIEGTCTGSSSSPEHEMSQSGFSAATTSNGPLLSTVGSYDWAG 376
 QY 361 GVFLYTSKESKSTFINMTTRVDSNDNDAYLGAAAIILNRRVQSLVGLGAPRYQHIGLVAMFR 420
 DB 377 GVFLYTSKESKSTFINMTTRVDSNDNDAYLGAAAIILNRRVQSLVGLGAPRYQHIGLVAMFR 436
 QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYHYEQTRGGQSVCP 480
 DB 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYHYEQTRGGQSVCP 496
 QY 481 PRGQARWQCDAYLYGSGQGFWRFGAALTVLGDVNGDKLTDVAIGAPGSEDNRGAVLYF 540
 DB 497 PRGQARWQCDAYLYGSGQGFWRFGAALTVLGDVNGDKLTDVAIGAPGSEDNRGAVLYF 556
 QY 541 HGTSGSGISPHSORIAGSKSLPRLOYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
 DB 557 HGTSGSGISPHSORIAGSKSLPRLOYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
 QY 601 PVLRVKAIMFNPFEVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 660
 DB 617 PVLRVKAIMFNPFEVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 676
 QY 661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPVLRLNF 720
 DB 677 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPVLRLNF 736
 QY 721 SLVGTPLSAFGNLRPLAEDAQRLLFTALFFPKKNCNDNICQDDLSITFFSPMSLDCLVVG 780
 DB 737 SLVGTPLSAFGNLRPLAEDAQRLLFTALFFPKKNCNDNICQDDLSITFFSPMSLDCLVVG 796
 QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYKVKSTLQNRQSRWRLACESASSTEV 840
 DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYKVKSTLQNRQSRWRLACESASSTEV 856
 QY 841 SGALKSTSCSINHPIFENSEVTENITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900
 DB 857 SGALKSTSCSINHPIFENSEVTENITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 916
 QY 901 QLELPVYAVTVMTVSHGVSTKYNFTASENTSRVMOHQVQVSNLQSRSLPISLVFLVPV 960
 DB 917 QLELPVYAVTVMTVSHGVSTKYNFTASENTSRVMOHQVQVSNLQSRSLPISLVFLVPV 976
 QY 961 RLNQTIVDRPQVTFSENLSTCHTKERLPSSHDFLAELRKAPVNVNCIAVQRIQCDIP 1020
 DB 977 RLNQTIVDRPQVTFSENLSTCHTKERLPSSHDFLAELRKAPVNVNCIAVQRIQCDIP 1036
 QY 1021 PFGIQESFNATLKGMLSPDWYIKTSHNHLIVSTAELFNDVSFTLLPGQGFVRSQTET 1080
 DB 1037 PFGIQESFNATLKGMLSPDWYIKTSHNHLIVSTAELFNDVSFTLLPGQGFVRSQTET 1096
 QY 1081 KVEPPEVNPPLIVGSSVGGILLIILALITAALYKLGEFKROYKDMMSGGPPGAEPO 1137
 DB 1097 KVEPPEVNPPLIVGSSVGGILLIILALITAALYKLGEFKROYKDMMSGGPPGAEPO 1153
 RESULT 2
 S00551
 leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
 N:Alternate names: complement-3 receptor alpha chain
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 22-Oct-1999
 R:Accession: S00551; I59078
 R:Pyrela, R.
 EMBO J. 7, 1371-1378, 1988
 A>Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the i
 A:Reference number: S00551; MUID:88312584; PMID:3044779
 A:Accession: S00551
 A:Molecule type: DNA
 A:Residues: 1-1153 <PYT>
 A:Cross-references: EMBL:X07640; NID:952982; PIDN:CAA30479.1; PID:952983
 A>Note: the authors translated the codon CAC for residue 569 as Gln

R;Sastre, L.; Roman, J.M.; Teplow, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A;Title: A partial genomic DNA clone for the alpha subunit of the mouse complement receptor
A;Reference number: I59078; MUID:86287312; PMID:2942940
A;Accession: I59078
A;Status: preliminary; translated from GB/EMBL/DBD
A;Molecule type: DNA
A;Residues: 11-44 <RES>
A;Cross-references: GB:M14293; NID:9198993; PIDN:AAA39484.1; PID:G554193
C;Genetics:
A;Gene: Mac-1
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homology
C;Keywords: cell adhesion; glycoprotein; transmembrane protein
F;1-16/Domain: signal sequence status predicted <SIG>
F;17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 76.2%; Score 4476; DB 2; Length 1153;
Best Local Similarity 74.3%; Pred. No. 1.2e-301;
Matches 845; Conservative 142; Mismatches 149; Indels 2; Gaps 2;

Qy 1 FNLDTENAMTFQENARGQSQVYVQVQGSVVVVGAPQEIIVAAANQGSLYQCDYSTGSCPEI 60
Db 17 FNLDTEHPMTFQENAKGFGQNVVQLGTSVVVAAPOEAKAVNQTGALYQCDYSTSRCHPI 76

Qy 61 RLQVPVAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTVVKGLCELFGSNLRQOPQK 120
Db 77 PLQVPPEAVNMSLGLSLAVSTVPOQLLAGPTVHONCXENTVYNGLCYLFSGNLRLPQQ 136

Qy 121 FPEALRGCPQSDIDAFIDGSGSIPIHDFRAMEFVSTWELKSKSLFSLMYSEEF 180
Db 137 FPEALRGCPQSDIDVFLIDGSGSINNIDFQWKEFVSTWELKSKSLFSLMYSEEF 196

Qy 181 RHFTKFEQNNPRSLVKPTOLLGRTHATGVRKVRRELLNTNGARKNAFKILLVI 240
Db 197 RHFTFNDPKRNPSPRSHVSPKQLNGRTKTASGIRKVRRELFHKTNGARENAKILLVI 256

Qy 241 TQGEKFGDPLGYEDVTPEDRGGVRYVIGVGDAPKRSKROELNTIASKPRDRHVFQVN 300
Db 257 TQGEKFGDPLGYEDVTPEDRAGVRYVIGVGNVFNKPSRRELDTIASKPAGEHVFQD 316

Qy 301 NFEALNTIQNLRKIFAIEGTQTSSEFHEMSQEGFSAITNGPILSTVGYDWAG 360
Db 317 NFEALNTIQNLRKIFAIEGTQTSSEFHEMSQEGFSAITNGPILSGVDFDWAG 376

Qy 361 GYFLYTSKESFTINATRVDSQNDNAYLCYAAITLNRVQSLVGLGAPYQHIGLVAMFR 420
Db 377 GAFLYTSKDKVTPINTIRVDSQNDNAYLCYAAITLNRVQSLVGLGAPYQHIGLVAMFR 436

Qy 421 QNTGWESNANVKGITGAYFGASLCSVDVDSNGSTDVLVGLGAPHYEQTRGGQSVCP 480
Db 437 ENFGTWEPHTSIKSGIYFGASLCSVDVDSNGSTDVLVGLGAPHYEQTRGGQSVCP 496

Qy 481 PRQORARQCDVLYGEOQPGWRCGALTVLGDVNGDKLTDAVAGPGEEDNRCGAVLYP 540
Db 497 PRG-RARWOCEALLHGDQHPMRFGAALTVDVNGDKLTDAVAGPGEEDNRCGAVLYP 555

Qy 541 HGTSGSGISPSHSORAGSKLSPRLQYFGQSLGGQDLTMDGLVLTGCAQGHVILLRSQ 600
Db 556 YGASIASLASHSRHIIGAFHPSFGLQYFGQSLGGQDLTMDGLVLTGCAQGHVILLRSQ 615

Qy 601 PVLVRKAIMENPREVARNVFEQNDQVKGKAGVRVCLVHOKSTRDLREGQIQSVTT 660
Db 616 PVLRLAETMEFSPKVARSVFAQEQVLYKNKDAGRVCLVRKNTKDLRGGDIQSTVT 675

Qy 661 YDLALDSGRPHSAVFNETKNSRTROTQVLGLTQTCETLKLQLPNCIBDPVPSVILRLNF 720
Db 676 YDLALDPVRSIRIAPFDETKNTNTRTQVPLGKQKCETLKLILPDCVDVDSVPIILRLNY 735

Qy 721 SLVGTFLSAFENLRPLVAEDAQLFTALPFEPKQGNNDNICQDGLSITFSMSLDCLVWG 780
Db 736 TLVGEPLRSFENLRPLVLAEDAQRFFTAAMPFEKQGNDSICQDGLSITFSMSLDCLVWG 795

Qy 781 GPREFNVTVTRNDGSDGSHVYRQVTFPPPLDLSTRKVTSTLQVRSQSRWRL-ACESASSTE 839
Db 796 GPQDFNMSVTLRNDGSDGSHVYRQVTFPPPLDLSTRKVTSTLQVRSQSRWRL-ACESASSTE 855

Qy 840 VSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANKVTSNNMPTVKTE 899
Db 856 GHGALKSTTWNINHPFPANSEVTFNITFDVDSKASLGNKLLKANKVTSNNMPTVKTE 915

Qy 900 FOELPVPKAYVMVTVSHGVSTKYLNFASNTSRVNHQYQVSNLQORSPLPISLVLPV 959
Db 916 FOELPVPKAYVMVTVSHGVSTKYLNFASNTSRVNHQYQVSNLQORSPLPISLVLPV 975

Qy 960 VRLNQTVIDRPPQVTSSENLSSCTCHTERLPSSDPLAELRKAPVNVCSSTAVCORICDI 1019
Db 976 VQINNVTVDHPQVIFSQNLSSACHTEQKSPFNSFRDQLERFVLNCSVAVCKRIQCDL 1035

Qy 1020 PFGIQEENFATLKNLSFDWYINTKSNHLLIVSTABILFNDSTVFTLLPGGAFVRSQTE 1079
Db 1036 PSEPTQIEFNVTLLKGNLSFDWYINTKSNHLLIVSTABILFNDSTVFTLLPGGAFVRSQTE 1095

Qy 1080 TKVEPFPVPLPLIVGSSVGGILLALITAAALYKLGFEKQYKDMSEGPPCAEPQ 1137
Db 1096 TKVEPFPVPLPLIVGSSVGGILLALITAGLYKLGFEKQYKDMSEGPPCAEPQ 1153

RESULT 3
RWHLIC
cell surface glycoprotein CD11c precursor - human
N;Alternate names: leukocyte adhesion receptor p150,95 alpha chain
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 22-Jun-1999
C;Accession: A36584; A35543; S00864
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 12750-12751, 1990
A;Reference number: A36584
A;Contents: erratum
A;Accession: A36584
A;Molecule type: DNA
A;Residues: 1-1163 <COR>
A;Note: This revision to the sequence from reference A35543 includes the carboxyl end
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A;Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule
A;Reference number: A35543; MUID:90153906; PMID:2303426
A;Accession: A35543
A;Molecule type: DNA
A;Residues: 1-834 <CO2>
A;Note: This sequence has been revised in reference A36584
R;Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A;Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A;Reference number: S00864; MUID:88166645; PMID:3327687
A;Accession: S00864
A;Molecule type: mRNA
A;Residues: 1-755, 'L', 757-1163 <CO3>
A;Cross-references: GB:M61695; EMBL:Y00093; NID:9487829; PIDN:AAA59180.1; PID:9487830
A;Note: part of this sequence was confirmed by protein sequencing
C;Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on m
C;Genetics:
A;Gene: GDB:ITGAX; CD11C
A;Cross-references: GDB:119758; OMIM:151510
A;Map position: 16p11.2-16p11.2
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C;Keywords: calcium, cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;
F;1-19/Domain: signal sequence status predicted <SIG>
F;20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>
F;20-1107/Domain: extracellular #status predicted <EXT>
F;149-319/Domain: von Willebrand factor type A repeat homology <VWA4>
F;1108-1133/Domain: transmembrane #status predicted <TM>
F;1134-1163/Domain: intracellular #status predicted <INT>
F;61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pr

A:Accession: A45226
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1151 <BRI>
A:Experimental source: hepatoblastoma cell line HepG2
A:Note: sequence extracted from NCBI backbone (NCBIP:124326)
F:142-317/Domain: von Willebrand factor type A repeat homology <VMA1>

Query Match 18.8%; Score 1102.5; DB 2; Length 1151;
Best Local Similarity 27.8%; Pred. No. 9.3e-66;
Matches 344; Conservative 211; Mismatches 487; Indels 195; Gaps 44;

Qy 1 FNLDENAMTPOENARG-FGQSVVQL---QGSRVVVGAPQEIIVAAKORGSLYQCDYGTGS 56
Db 1 FNVDKSMITSGPVEDMFGYVQYENEBGKWLIGSLPLVGQPKNRTGDKYKFCVGRGE 60

Qy 57 CEP-IRLOVPEA-----VMSGLSLAATISPPQLACGPTVHQTCSNTYKGL 106
Db 61 SLPCVKLDLPVNTSIPNVTEYKNTFGSTL-VTNPNGGFLACGLYAYRCGHLYTGI 119

Qy 107 CFLFGSNLRQOPQKFPPEALRCPOEDSDIAPLDGSGSIIPHDPRRMEKFVSTVMEQLK- 165
Db 120 CSDVSPFTQVNSIAP-VQECSTQ-LDIVIVLQSGNSIYPWD--SVTAFNLLDKEMDI 174

Qy 166 -KSTLFLSMOYSEEFRIHFTFKFQNNPNRSLVKPITQLLGR-THTATGVRKVIKRELL 223
Db 175 GPKQTOGVIGYGENVTHEFNANKYSSTEEVLAAKKIVQGRGRTWTALGTDTRKEAF 234

Qy 224 NITGARKNAFKILVITDGEKFDPLGYEDVIPEADREGVIRVIVGVDAFR-----SE 278
Db 235 TEARGARGVKVWIVITDGEH-DNRLKVKVIOCEDENIQRESIALGSGYKNGNSTE 293

Qy 279 KSRQELNTIASKPRDRHVQNNFEALKTIQNLREKIPAEFTGTQSGSSFEHEMSQEG 338
Db 294 KXVEBIKSIASEPTEKHPFNVSDELAVITVKTGERIFALEATADQSAASFEMENQ 353

Qy 339 FSAITNSGPIVLSVGSYDAGVGLYTSKE-----KSTP-INNTVDSMDNDVLYGAA 392
Db 354 FSAHYSDWMLGAVGADNGTVGMQKASQIIIPRNTFNVESTKKNELP-ASLTGTV 412

Qy 393 AIIENLRVQSL-VLGAPRYOHIGLVAMPRQNTGMWESNANVKGITQIGAYFGASLCSDVD 451
Db 413 NSATASSGDVLYIAGQPRYNTGQVILYRMDGNKIKILQSLGSGIYSGYFISLTITDID 472

Qy 452 SNGSTDLVLGAPHY-----YEQTR-GQGVSVCLPQGRARWQCDVLY 495
Db 473 KDSNTDILLVAGAPYMGTEKEEGKVYVYALNTRFQMSLEPIKOTCCSSRQHNCTT 532

Qy 496 GEOQOPWG-RFCALTLVLDVNGDKLTDVATGAPCEDNMGAVLYFHTGSGSGISPSHSQ 554
Db 533 ENKHEPCGARTGATAAVKDLNLDGFDNDIVIGAPLEDHGGAVIYHG-SGKTIRKEYAQ 591

Qy 555 RIAGKLSPLRQYFQSGLSGGODLTMDGLVDLTGVAQGHVLLRSQPLVLRKAIWENPR 614
Db 592 RIPSQGGDKTLKFFQSGIHGMDLNGDGLTDVTIGLGAALFWSRDVAVVVKVTMNPEN 651

Qy 615 EVARNVFCNDQVVKGEAG-EVRVCLHVQ-KSTRDELREGQIQSVVYDIALDSGRPH 671
Db 652 KVNIOKKNCH---MEGKETVCINATVCEVKLSKEDTIYRADLQ-----YRVLDSLRQI 704

Qy 672 SRVAFNET-----KNSTRRTQVIGLQTCETLKLQLPNCIEDPVPVILRLNPSLVGT 725
Db 705 SRSFSGTQERKQVQNRNIVRKSEC-----TKHSFYMLDKHDFQDSVR---ITLDFNLT-D 755

Qy 726 PLSAFGNLRPLVLAEDAQLFTALPFPEKNCNDNIQDGLDITFSFMSLCLVVGPRE- 784
Db 756 PENG-----PVLDDSLPNSVHEIYFPKXDCGKKEKISDLSLHVAETEKULLIYRSQDK 810

Qy 785 FNVTVVRNDEGDSYRTQVTFPPFDLSYRVKSTIQNORSQSWELACESASSTEVSGAL 844
Db 811 FNVSLTVKNTKDSANTRTIVHYSNLFVSGIEALQKD-----SCSEN----- 853

Qy 845 KSTCSINHPFPENSEYTFNITFDVDSKASLGN-KLLKKNVTSENMPRTNKTQLE 903

Db 854 ENITCKVGVFLRRGEMVTFKILFQNTSYLMBNVIYLSATSDSEPPETLSDNVNIS 913
Qy 904 LPVKYAVYVWVSHGVSTKYLNFATSENTRYMOHQYQVSN-----LQGRS-----L 950
Db 914 IPKYEVGLQFYS-SASEYHISIAANETVPEVINSTEDIGNEINIFYLRKSGSPMPEL 972
Qy 951 PISLVE-----LVPRLNQTVIMDRPQVTFSENLSSTCTHKE-----RLPS 991
Db 973 KLSISPPNMTSGYPVLYPTGLSS-----SENANCRPHIFEDPFSINGKQWTT 1021
Qy 992 HSDPLAELKAPVWVCSIAVCQRIQCDIFFFGIQE-----EFNATLK 1033
Db 1022 STD---HLKRGTTLDCTCKFATITCNLTSSDISQVNSVLILMKPTTFIKSYFSSLNLFIR 1078
Qy 1034 GNLSFDWYIKTSHNHLIIVSTABILLNDVSFTLLPGCGAFVRSQTEKVEPFPVPLPL 1093
Db 1079 GEI-----RSENASVLSSSN-----QKRELAIQISKDGLGKRVPL 1114

Qy 1094 --IVGSSVGGLLALLALITAAALYKLGPKQYKXDMSE 1128
Db 1115 WVILLSAPAGLILLMELLILALMKIGFPKPLKKMEK 1151

RESULT 8

I45914
Integrin alpha 2 subunit - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 15-Sep-2003
C:Accession: I45914
R:Kamata, T.; Puzon, W.; Takada, Y.
J. Biol. Chem. 269, 9659-9663, 1994
A:Title: Identification of putative ligand binding sites within the I-domain of integrin
A:Reference number: A54402; PMID:94193647; PMID:7511592
A:Accession: I45914
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1170 <KAM>
A:Cross-references: GB:L25886; NID:9439695; PIDN:BA859255.1; PID:9439696
F:161-336/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match 18.5%; Score 1084; DB 2; Length 1170;
Best Local Similarity 27.7%; Pred. No. 1.8e-66;
Matches 336; Conservative 216; Mismatches 495; Indels 168; Gaps 47;

Qy 1 FNLDENAMTQ-ENARGGQSVVQL---QGSRVVVGAPQEIIVAAKORGSLYQCDY- 54
Db 19 YNVGLPKAKIFSGPSSEQFGYAVQGFNPNGNMLLVGSPWSPGPKRMGVDVYKCPVDLST 78

Qy 55 GSCPEIRLQ-----VPVEAVNWSLGLSLAATISPPQLACGPTVHQTCSNTYKGLC 107
Db 79 TTCEKLNQTSMSNVTEMKTNWSLGLTLTRNVGTGGFLTCGPLMAQQCGSQVYTTGVC 138

Qy 108 FLFGSNLRQOPQKFPPEALRCPOEDSDIAPLDGSGSIIPHDPRRMEKFVSTVMEQLK- 165
Db 139 SDVSPDP-QLTSTFAPAVQTCF-SFIDVVVVCESNSIYPWD--AVKGFLEKFPVQGLDIG 194

Qy 166 KSKTLFSLMOYSEEFRIHFTFKFQNNPNRSLVKPITQLL-----GRHTATGVRKVIKRE 221
Db 195 PTKTQKGLIYANNPRVFNLTNFKSD---EMKATSTQFYCGDLTNTFKAIQVARDT 251

Qy 222 LAMITNARKNAFKILVITDGEKFDPLGYEDVIPEADREGVIRVIVG-----GDAPR 276
Db 252 AYSTAAGRGPRGATKVMVWVTDGESH-DGSKLKAVIDQCNKDNILRFGIAGLVGLYNRLD 310

Qy 277 SEKSRQELNTIASKPRDRHVQNNFEALKTIQNLREKIPAEFTGTQSGSSFEHEMSQ 336
Db 311 TKOLIKIKAIASIPTRHFFNVSDDEADLLEKAGTIGEIQIFSIETGVQ-GDNFQEMSQ 369

Qy 337 EGFSAAIT--SNGPLISTVGSYDAGVGLYTSKESKSTFTNMT--RVSDSMN-DAYLGYA 391
Db 370 VGSFSAEYSPQNTILMLCAVGAYDWSGTVQKTPHGLIFSKQAFEQIQLDRNHSYLGYS 429


```

QY 956 FLV-----PVLNQTWIDRRQVTF-SENLS 980
D5 1VTTHIPQYTKENPLLYLTGQTQDAGDISCTAEINPLKLPHFA-----DSVSFKNENR 1040
QY 981 STCHTKERLPSHDFLAELKAPVWNGSIAVCQIQDIPFGIOBEPNATLKNLSFDW 1040
D5 1041 ----HTKE-----LDCRTTSCSNITCWLKDLHKABYFINVTVRWNR 1080
QY 1041 YIKTSHNHLIVSTAELFNDVSFTLAPQCGAFVRSQTETKVEPFFVNPPLIVGSSVG 1100
D5 1081 FAASTFQVQLTAAAEIDTHNPQLFVIEENAVTIPLMIMKTEKAEVPT--GVIIGSIIA 1138
QY 1101 GLILLALITAAALYKLGFFKROYK 1125
D5 1139 GILLALITAAALYKLGFFKROYK 1163

RESULT 10
A:33998
Integrin alpha-2 chain precursor - human
N:Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 ch
C:Species: Homo sapiens (man)
C:Date: 30-Mar-1990 #sequence revision 18-Sep-1992 #text_change 15-Sep-2003
C:Accession: A33998; B56793; A53117
R:Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A:Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet
A:Reference number: A33998; MUID:89308879; PMID:2545729
A:Accession: A33998
A:Molecule type: mRNA
A:Residues: 1-1181 <MAX>
A:Cross-references: GB:X17033; MID:g33906; PIDN:CAA34894.1; PID:g33907
A:Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
R:Catimel, B.; Parentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A:Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIC*, GPiIa an
A:Reference number: A56793; MUID:92061944; PMID:1953640
A:Accession: B56793
A:Molecule type: protein
A:Residues: 30-43 <CAT>
A:Experimental source: platelet
R:Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A:Title: The human alpha-2 integrin gene promoter. Identification of positive and negati
A:Reference number: A53117; MUID:94103255; PMID:8276836
A:Accession: A53117
A:Molecule type: DNA
A:Residues: 1-16, 'v', 18-21 <ZUT>
A:Cross-references: GB:L24121; MID:g400342; PIDN:AAA16619.2; PID:g4583535
A:Note: authors translated the codon GTA for residue 17 as Leu
C:Genetics:
A:Gene: GDB:ITGA2; CD49B
A:Cross-references: GDB:128031; OMIM:192974
A:Map position: 5q11.1-5q11.2
C:Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-1133/Domain: extracellular #status predicted <EXT>
F:172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1134-1154/Domain: transmembrane #status predicted <TM>
F:1155-1181/Domain: intracellular #status predicted <CYT>
F:1105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 18.2%; Score 1069; DB 2; Length 1181;
Best Local Similarity 27.0%; Pred. No. 2e-65;
Matches 331; Conservative 212; Mismatches 494; Indels 188; Gaps 43;

QY 1 ENLDTENAMTQ-ENARFGGSVQL---QSSRVVVGAPQIIVAANQSGSLYQC--DYST 54
D5 30 YNVGLPEAKIFGSPSSQFQYAVQQFIPNKGWLLVGSFWSGPPENRGDVKPCVDLST 89
QY 55 GSCPEIRLQ-----VPVEAVNMSLGLSLAATSPQQLACGPTVHQTCSNTYVKGLC 107
D5 90 ATCEKLNLTQSTSIPNVTEMKTNMSLGLILTRNMGTCGLTQGLPWAQCCNQYVYTGVC 149

```

```

QY 108 FLFGSNLRQOPQKPPALRGCCPOEDSIAFLIDGSGIIPHDPFRMKBFVSTVMEQLK-- 165
D5 150 SDISPDV-QLSASFSEATQPCSL-IDVVVCCDESNSIYPWD--AVKNFLKFKVQGLDIG 205
QY 166 KSTLPSLAWOYSEBFIHFTFKFQNNPNSRLVAKITQLLG-RTHTATGVKRVIRELN 224
D5 206 FTKTQVGLIQYANNPVRVFNLTNYTKKEMIVATSTQSYGGDLTNFQIGAIQARKAYS 265
QY 225 ITNGARKNAFKILIVITDGEKFDGPDIGYEDVITPEADREBVGIRYVIGV-----GDAPRSEK 279
D5 266 AASGGRSRTKVMVVVTDGESH-DGSMLKAVIDQCNDHNDILRPGIAVLGYLNRNALDTKN 324
QY 280 SROELNTIASKPRDRHVQNNFEALKTQNLREKI PAIEGTQTGSSSSFEHMSQEGF 339
D5 325 LIKEIKATIASIPTEYFFNVSDAALKEKAGTGEQIFSEIEGTVOG-GDNFQEMSGVGF 383
QY 340 SAATITNGP--LLSTVGSYDWAGGVFLYTSKSKSTFINMT--RVDSDMN-DAYLGYAAAI 394
D5 384 SADYSSQNDILMCAVAGFMSGTIVQKTSKSHGLIFPKQAFDQILODRNHSSVLYGVAA 443
QY 395 ILNRVQSILVGLAPRYOHIGLVAMFRONTGMWESNANV-----KGTQICAYFGASLCVS 448
D5 444 ISTCESTHFVAGAPRANYTQIIVLYSVN-----ENGNITVIQAHRGDQIGSYFGSLCSV 498
QY 449 DVDSNGSTDLVLGAPHYEYQTR--CGOVSVCPLPRGORARMOCDVLYGEOQOPWGRPG 506
D5 499 DVDKDTITDVLVGRPMWMSDLKKEGRVYLFPIKIGILQCHQ---FLEGPSGIENTRPG 555
QY 507 AALTVLGDVNGDKLTDVAIGAPCEEDNRGAVLYPHGTSGSGISPSHSQRTAGS--KLSPR 564
D5 556 SAJAAALSDINMDFNDVIVGSPLENQSGAVIYNGHQGT-IRTKYSOKILGSDGAFRSH 614
QY 565 LQYFGSLSGGDLTDWGLVLTGAGQHVLLRSQPLVLRKAIMRPNPREVARNVPECN 624
D5 615 LQYFGESLQYGLNGDSITDVSIGAPGVQVQLMSQSIADVAIEASTPBEKI--TLVNKN 672
QY 625 DQVVGKGEAGEVRVCLHVQKSTRDLRREGQIQSWVTYDIALD---SGRPHSRAVFNETK 680
D5 673 AQII-----LKLCP-----SAKERPTKQNNVAIVNITLDADGFSRVTSGRLPKENN 721
QY 681 NSTRTOVQLGLTQTC--ETLKLQLNCEIDPSPVILRNPSILVGTPLSFAFGNLRVLA 738
D5 722 ERLCKNMVWVQACSPHEIYIQEPS--DVVNSLDLRVDISLENGTGS-----PALE 772
QY 739 EDAQRLFTALFPPEKNCNDNIQDDLSITF-----SFMSLDCLVGVGPRFNVTVVRND 794
D5 773 AYSETAKVFSIPFKOCGEDGLCISDLVDVQIPAAQEQPFIVSNQKELTFSVLKNK 832
QY 795 GEDSYFQVTFPPLDLVYKVKSTLQNRQSRWRLACESAST-EVSGALKSTSCSINH 853
D5 833 REGAYNTGIIVDFSENLF-----ASFSLPVDGTGTEVTCQVAASQKSVACDVGY 880
QY 854 PIPENSEVTNLTDFDVSKASLGNKLLKANVTSENNMPRTNKTETFOLELPVKIAYMV 913
D5 881 PALKREGQVTFITNDFNLQ-NLQNASLSFQALSQEQENKADNLVNLKIPLYDAEI- 938
QY 914 VTSHGVSTKYLNFTASENTSRVWQHGYQVSNLQOR-----SLPISLVFLV----- 958
D5 939 ---HLRSTNINFEIISDGNVPSIVHSFEDVGPKEIFSLKVTGTGSPVSVMAVTHIPQ 995
QY 959 -----PVLNQTWIDRRQVTF-SENLSSTCHTKER 988
D5 996 YTEKKNPLMYLTGVQTDKAGDISCNADINPLKIGQT-----SSVSFPAKENFR---HTKE- 1047
QY 989 LPGHSDFLAELKAPVWVNCISIAVCQIQDIPFGIOBEPNATLKNLSFDWIKTSHNH 1048
D5 1048 -----LNCRTASCNVTCWLDVHKMGVEYFVNVTVTRWNGTFEASTFQT 1091
QY 1049 LLIVSTRABI-LFNDSTVFTLLPGQAFVRSQTEKTVKPFVEPNP-----LP--LIVGSSVG 1100
D5 1092 VQLTAAABINTPNPIYVI-----BDNTVTIPLMIMKPEKAEVPTGVLIIGSIIA 1141

```



```
QY 156 FVSTVMEQLKSKTFLSLMOYSEEFRIHFTKFEQNNPNRSLVKPITQLLGRHTATGV 215
Db 186 -----RWAPCVKDYT----- 195
QY 216 RKTRELLNITNGARKAKFLIVITGEGKEDPLGVEDVPEADREGVIRYVIGVDAP 275
Db 196 -----RKPE----- 200
QY 276 RSEKSRQELNTIASKPPRHHVFOVNNFEALNTIQNLREKIFAIEGTQTGSSSFEHMS 335
Db 201 -----NFAS-----C 205
QY 336 QEGESAITSNGPILLSTVGSVDNAGGVPLY--TSKEKSTFINMTRVDSMDNDAYLGYA-- 391
Db 206 QAGISSFYTDLIYMGAPGSYMTGTTFVYNNITNQYKAFVD--RQNKVFGSGLSVSG 263
QY 392 AAIIILNRVQSLVIGAPRYOHIGLVAMFRONTGHWESNANV---KGTQIGAYFGASLCS 447
Db 264 AGHFRSHPTTEVVGAPQHQEQIKAYIF---SIDENELNIYEMKKGKLSYFGASVCA 319
QY 448 YVDNSGSTDVLVLGAPHYVQTRGGQVSVCLPRGQRA--RWQCDVLYGEGQGFWRGFG 506
Db 320 VDLNADGFSDL--LVGAPMQSTIREGRVFVY--INSGMGAVMVENRVLVGSDDKYA--ARFG 376
QY 507 AALTIVLGDVNGDKLTDVAIGAPGEDNREGAVYLFHGTSGSISPSHSORLAGSKLSRLQ 566
Db 377 ESIANLGDIDNDGPDIAIGAPQDDDLRGAVIITNGRV--DGISSTYSQRIEQQISLSLR 435
QY 567 YFGQSLSGQDLTMDGLVDLTVGA--QGHVILLRSQPLRVKAIMENPREVARNVFEEN 624
Db 436 MFGQSIQSQIDADNNGYVDVAVAGPQSDSAVLLATRPVIVASLS--HPESVNETKEDCT 494
QY 625 DOVVKGEAGEVRLVCHVOKSTRDLREGQIQSVVYDLDLDSGR---PHSEAVF--NET 679
Db 495 -----ENGLPSVCMHLTLCFSYKGEVPGYIVLFYNVSLDVHKKAESPRFTFFNGT 547
QY 680 KNSTRROTQVLGLTQFCBTLKQLPNCIEDPVSPTVLRNLSL-----VGTPLSAPGNLR 734
Db 548 SDVITGSRVSSGSEKCTHOAFKRVORDLITPIHVEATVHLGHVITKNTSEEFPLQ 607
QY 735 PVLAEDQR--LFTALPPEKKNCGDNICQDLSITFS-----FMSLDCLVVGQPRENV 787
Db 608 PILQOKKEQVIRKMINFAFCAYEN--CSADLOVSARKVGLPKPYENKTYLAVGSMKTMIL 666
QY 788 TVTVNDEGDSVKTQVTFPPFLDLSYKVSFLQORSORSLACESASSTEVSGALKST 847
Db 667 NVSLFNAGDADYETTLNVQLPTGLYFKILDLBK-----QINCE-----VTIESSGVK-L 716
QY 848 SCSSNHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSEN--NMPTKTKTFQLELP 905
Db 717 ACSLGYIVVDRLSIDISFLDVSLSRAHEDLSISVHASCENEGELDQVRDNRVLTIP 776
QY 906 VKYAYVMVTSQV--STKYLNPASENTSRVMQHQ-----YQVSNLQORSLP--ISLVFL 957
Db 777 LREYEV--MLTVHGLVNPFSFVVGSEENEPETCAEKLNLTFTHVINTQISVAPNVSKIM 834
QY 958 VEVRLNQTVINDRP--QVTFSENLSSTCKTE-----RLPHSDPFLAELR 1000
Db 835 VP---NSFLPQDDKLFNVLVDVTTTGQCHFKYHGRECTFAQCKGIAGTLTDIVKFLSKTD 891
QY 1001 KAPVNVCSIA--VCQRIQCDIPFGIQEFPNATLKNLSFMYIKTSHNLLIYSTAEIL 1058
Db 892 KR--LLYCKMADQHCFLDPLCN---FGKMSG-----KEASVHTQLGERPSIL 933
QY 1059 FNDVSFTLLPGGQFVRSOTETKVPFFVFPNP----- 1090
Db 934 ENDETSSL-----KPEIKATAFPEPKPVIENLKNDENVAHVLEGLHQRKHEF 983
QY 1091 -LPLIVGSSVGLLLALITAAIYKLGFFKQYKDMSE 1128
Db 984 TIIITISILLGLIIVLLISCVMKAGFFKQYKSILOE 1022
```

```
RESULT 13
S06046
Integrin alpha-4 chain precursor - human
N;Alternate names: CD49d; very late antigen-4 alpha chain; VLA-4 alpha chain
C;Species: Homo sapiens (man)
C;Date: 01-Dec-1989 #sequence revision 01-Dec-1989 #text_change 24-Sep-1999
C;Accession: S06046; A39355; D28018
R;Takada, Y.; Ellices, M.J.; Crouse, C.; Hemler, M.E.
EMBO J. 8, 1361-1368, 1989
A;Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other inte
A;Reference number: S06046; MUID:89356603; PMID:2788572
A;Accession: S06046
A;Molecule type: mRNA
A;Residues: 1-1038 <TAK>
A;Cross-references: GB:X15356; EMBL:X15356; NID:G33945; PIDN:CAA34852.1; PID:G33946
R;Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.
Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991
A;Title: Characterization of the alpha-4 integrin gene promoter.
A;Reference number: A39355; MUID:91239513; PMID:2034655
A;Accession: A39355
A;Molecule type: DNA
A;Residues: 1-93 <ROS>
A;Cross-references: GB:M62841
R;Takada, Y.; Strominger, J.L.; Hemler, M.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987
A;Title: The very late antigen family of heterodimers is part of a superfamily of molec
A;Reference number: A94151; MUID:87264112; PMID:3033641
A;Accession: D28018
A;Molecule type: protein
A;Residues: 40-50, 'E', 52-53 <TA2>
C;Genetics:
A;Gene: GDB:ITGA4; CD49D
A;Cross-references: GDB:I28032; OMIM:192975
A;Map position: 2q31-2q32
C;Superfamily: integrin alpha-4 chain
C;Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; transme
F;1-39/Domain: signal sequence #status predicted <SIG>
F;40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>

Query Match 10.8%; Score 633; DB 2; Length 1036;
Best Local Similarity 25.2%; Pred. No. 2.8e-35;
Matches 248; Conservative 155; Mismatches 370; Indels 212; Gaps 39;

QY 272 GDAPRSEKSRQELNTIASKPPRD-----FAEGTQSGSSFEHMSQEGFSAITNSGPLLSTVGSY 356
Db 121 GKTCLERDQNLGVTLSRQKENGSIIVTCGHRWKNIIFYIKNKENKLPCTGCGYVPPDLRT 180
QY 311 QLEKEL-----FAEGTQSGSSFEHMSQEGFSAITNSGPLLSTVGSY 356
Db 181 ELSEKRIAPCTQDVVKKFGENPA-----SCQAGISSFTKDLIVMGAPGSS 225
QY 357 DWAGGVPLY---TSKEKSTFINMTRVDSMDNDAYLGYA--AAIILNRVQSLVIGAPRYQ 411
Db 226 YWTGSLPVYNNITNKYKAFLDKQNVKF---GSLGYSVGAGHFRSQHTTEVVGAPQHE 282
QY 412 HIGLVAMFRONTGHWESNANV---KGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHY 467
Db 283 QIKKAYIF---SIDEKELNIHEMKKGKLSYFGASVCAVDLNADGFSDL--LVGAPMQS 337
QY 468 EQTRGGQVSVCLPRGQRAWQC--DAVLYGEGQGFWRGFGAALTVLGDVNGDKLTDVAIG 526
Db 338 TIREGRVFVY--INSGMGAVMVENRVLVGSDDKYA--ARFGESIVNLGDIQNDGPELVAIG 395
QY 527 APGEDNREGAVYLFHGTSGSISPSHSORLAGSKLSRLQYFGQSLSGQDLTMDGLVDL 586
Db 396 APQEDDLQGAIIYNGRA--DGISSTYSQRIEQQISLSMFGQSIQSDADNNGVVDV 454
QY 587 TVGA--QGHVILLRSQPLRVKAIMENPREVARNVFEENQVVKGEAGEVRLVCHVOK 644
Db 455 AVGAFRSDSAVLLATRPVIVASLS--HPESVNETKEDC-----VENGWPSVICDILT 506
QY 645 STRDLREGQIQSVVYDLDLDSGR-----PHSEAVFNETKNTSTRQTVLGLTQTCETL 699
```


A;Residues: 1-1041 <HER>
A;Cross-references: EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:ARC23572.1
A;Experimental source: developmental stage embryo
C;Function:
A;Description: binds laminin
C;Superfamily: integrin alpha-2b chain

Query Match 10.58; Score 614.5; DB 2; Length 1041;
Best Local Similarity 24.88; Pred. No. 5.4e-34;
Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

QY	336	QEGFSAITSGNP--LLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDS-----DKMDAYLG 389
DB	181	QAGFSGIIFSDNSALVNGAPGSGYVLOGQIYVQSLNRSV-VQATQESNTGTYSFDNSYRG 239
QY	390	YAAAI--ILNRVQSVLGCAPRYQHI-GLVAMPFQNTGMWESNANVKGTGTOIGAVFGASLC 446
DB	240	YSLALGDFNGDGVDQYVVGTPRAESLMGLVAIFDNLNQFN---QVMGTQIVAFEGYSVT 296
QY	447	SVYVDSNGSTDLVLIGAPHYEQTGGQVSVCLPRGORARWQCDVLYGEQ-----498
DB	297	VVDI--NNDTYDLLVGAFWYMDGPAIQ-----RWEGAVVYVLQNPDVGPGA 343
QY	499	-----GQPMWFGAALTIVLGVNGDKLTDVAIGAPGEEDNRGAVYLEHGTSGS 546
DB	344	SNRLSSTLIGGQTRSRFGSLIASIGDSNQDGFNDVAIGAPYEGDDAGAVYIYHG--SAN 402
QY	547	GISPSHSQRIAGKLS--PRLQYFGOSLGGQDLTMDGLDVLTVGAQ--GHVLLERSQPV 603
DB	403	GLKSTPAQVLTPSTLGHSGITTFGFSLOGGQDMENKYPDLLVGAESANTAVLIRTPVV 462
QY	604	RVKAIMFNPREVARNFECNDQVVVKEAGEVRVCLHVQKSTRDRRREGIQSV-----558
DB	463	SILDATLNTPE-----IGINLENKTYE-LADGTMVTSFIAMT 497
QY	659	-----VTYDLALDSG-RPHSRAVENETKNSTRQTQVGL-TOTCETLKIQ 702
DB	498	CFTYTGNYLPDHIDISYTVVDSGIIANRRAMFVNDMSBITKRRRAVSTQFCDFLEAY 557
QY	703	LPNCIEDPVSVPLRLNPLNPLSAPGN-----LRPVLAEADQRLFTALFPPEK 753
DB	558	VGNSTIEDKLTPIKVTLOYDL-----NNDSERLQPHLEILLIDMATMSTQTKQVSIQN 609
QY	754	NCGNNDICODLSITPFSFMSLDCLVVGPREFNVTVTVNRDGEDSYRTQVTFPELDLSY 813
DB	610	NCVN-NICIPDLDTVT-PNLFNIVIGQTELTDLVSLNRGEDAFQSSLSVYPIGLQF 667
QY	814	RVUSTLQORSQSRNLACESASTEVSCALKSTSCSINHPIFPEN-----SEVTNIT 867
DB	668	VRL-----ERKAMDPSVTCSESD-----LRITCDIGNPMVKNNILFGLTLSTFOVS 717
QY	868	FDVDS-----KASLGNKLLKANVTSENMPRTNKTPEQLPVKVATVMVTSHGVS TK 922
DB	718	GDKDSIEFYFKAESENS--EDPNTLENNELNMTVPVTVDTCLKLSASYPEIWMYSTQED 775
QY	923	YL-----NFTASENTSRVMOHQYV-----SNLQSRSLPIS-----LVFLVPV 960
DB	776	YVVPFPKXNASEADIGMEVM-HLYEVRNTGSGNAGEVSLNQWPKNEDGEYLFYLLGI 834
QY	961	RLNQTVIMDRPQ-----VTFSENLSSTCHTKERLPKSHSDFLAELKAPVV 1005
DB	835	MTEEGVTCQLTQKAPNEGKLEPSTKAKLSNSTQVSGRKRREPEVAELACTDN--VI 892
QY	1006	NCSIAVCQRIQCDIPFPGIQEENAT-----LKNLSFDWYIKTSHNHLIVSTAELP 1059
DB	893	YCASDSCVLINCTI-----DEINASKSKVVRILGRF--W-----ERTF 928
QY	1060	NDSVFTLLPGQGFVRSQETKVE--PFEVPNP-----LP-----1092
DB	929	QKAVSELTVPVQATIASAAVKTIPIYNIPLPRDFSDSTKASTLVTTTELVPVPTPIAM 988
QY	1093	-LIVGSSVGLLLALITALYKLGPFKR-----QYKDMMS--EGGPP 1132

DB 989 WIIWVSVLGGIILLIILGLWKCGFFERKXPGEEKYAPVASADKGGPP 1038
Search completed: June 7, 2004, 17:18:19
Job time : 25.559 secs

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:55:38 ; Search time 9.97194 Seconds
(without alignments)
5937.039 Million cell updates/sec

Title: US-09-902-481B-6
Perfect score: 5875
Sequence: 1 FNLDTENMTFQENARGFGQ.....FKRQYKDMSEGPPGAEFPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5852.5	92.6	1152	1 ITAM_HUMAN	P11215 homo sapien
2	4476	76.2	1153	1 ITAM_MOUSE	P05555 mus musculus
3	3473	59.1	1163	1 ITAX_HUMAN	P20702 homo sapien
4	3417	58.2	1162	1 ITAD_HUMAN	Q13349 homo sapien
5	1552.5	26.1	1170	1 ITAL_HUMAN	P20701 homo sapien
6	1533.5	26.1	1163	1 ITAL_MOUSE	P24063 mus musculus
7	1157.5	19.7	1167	1 ITAE_MOUSE	Q60677 mus musculus
8	1149	19.6	1179	1 ITAE_HUMAN	P38570 homo sapien
9	1102.5	18.8	1151	1 ITAI_HUMAN	P56199 homo sapien
10	1093.5	18.6	1189	1 ITAH_HUMAN	Q9UKX5 homo sapien
11	1084	18.5	1170	1 ITA2_BOVIN	P53710 bos taurus
12	1072	18.2	1178	1 ITA2_MOUSE	Q62469 mus musculus
13	1069	18.2	1181	1 ITA2_HUMAN	P17301 homo sapien
14	1065.5	18.1	1167	1 ITAG_HUMAN	O75578 homo sapien
15	1060	18.0	1180	1 ITAI_RAT	P18614 rattus norv
16	663	11.3	1039	1 ITA4_MOUSE	Q08651 mus musculus
17	633	10.8	1038	1 ITA4_HUMAN	P13612 homo sapien
18	630	10.7	1035	1 ITA9_HUMAN	Q13797 homo sapien
19	593.5	10.1	1032	1 ITA4_XENLA	Q91687 xenopus lae
20	571.5	9.7	1066	1 ITA3_CRISP	P17852 cricetidae
21	567.5	9.7	1053	1 ITA3_MOUSE	Q62470 mus musculus
22	555.5	9.5	1053	1 ITA5_MOUSE	P11688 mus musculus
23	546.5	9.3	1050	1 ITA5_XENLA	Q08274 xenopus lae
24	543.5	9.3	1034	1 ITAV_CHICK	P26008 gallus gall
25	538.5	9.2	1130	1 ITA6_HUMAN	P23229 homo sapien
26	535	9.1	1044	1 ITAV_MOUSE	P43406 mus musculus
27	532	9.1	1049	1 ITA5_HUMAN	P08648 homo sapien
28	531.5	9.0	1066	1 ITA3_HUMAN	P26006 homo sapien
29	530.5	9.0	1072	1 ITA6_CHICK	P26007 gallus gall
30	526	9.0	1048	1 ITAV_HUMAN	P06756 homo sapien
31	517	8.8	1044	1 ITA8_CHICK	P26009 gallus gall
32	512.5	8.7	1091	1 ITA6_MOUSE	Q61739 mus musculus
33	498	8.5	1179	1 ITA7_MOUSE	Q61738 mus musculus

34	494	8.4	1396	1 ITA2_DROME	P12080 drosophila
35	491.5	8.4	1146	1 ITAI_DROME	Q24247 drosophila
36	489.5	8.3	1039	1 ITAB_HUMAN	P08514 homo sapien
37	489	8.3	1025	1 ITAB_MOUSE	P53708 homo sapien
38	489	8.3	1033	1 ITAB_MOUSE	Q9QUM0 mus musculus
39	486	8.3	126	1 ITAM_CAVPO	P11578 cavia porce
40	472	8.0	1181	1 ITA7_HUMAN	Q13683 homo sapien
41	469.5	8.0	1106	1 ITA7_RAT	Q63258 rattus norv
42	466	7.9	1226	1 PAT2_CAEEL	P34446 caenorhabdi
43	445.5	7.6	1139	1 INAI_CAEEL	Q03600 caenorhabdi
44	423	7.2	1115	1 ITA3_DROME	O44386 drosophila
45	383	6.5	1000	1 ITA5_DROME	Q9W1M8 drosophila

ALIGNMENTS

RESULT 1
ITAM_HUMAN STANDARD; PRT; 1152 AA.
ID ITAM_HUMAN
AC P11215;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)
DE (Neutrophil adherence receptor).
GN ITGAM OR CR3A OR CD11B.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88315033; PubMed=2457584;
RA Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;
RT "The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B."; J. Biol. Chem. 263:12403-12411(1988).
RL [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=88190151; PubMed=2833753;
RX Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;
RA "Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol: chromosomal localization and homology to the alpha subunits of integrins."; Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).
RL [3]
RN SEQUENCE FROM N.A.
RP MEDLINE=88257215; PubMed=2454931;
RX Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;
RA "Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mol (complement receptor type 3)."; J. Cell Biol. 106:2153-2158(1988).
RL [4]
RN SEQUENCE FROM N.A.
RP MEDLINE=93123748; PubMed=8419480;
RX Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.P., Tenen D.G.;
RA "Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution."; J. Immunol. 150:480-490(1993).
RL [5]
RN SEQUENCE OF 9-1153 FROM N.A.
RP MEDLINE=89098893; PubMed=2563162;
RX Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;
RA "cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits."; Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).
RL [6]
RN

RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=92073318; PubMed=1683702;
RA Shelley C.S., Arnaut M.A.;
RT "The promoter of the CD11b gene directs myeloid-specific and
developmentally regulated expression.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7]
RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=92144986; PubMed=1346576;
RA Paul H.L., Rosmarin A.G., Tenen D.G.;
RT "Characterization of the myeloid-specific CD11b promoter.";
RL Blood 79:865-870(1992).
RN [8]
RP SEQUENCE OF 17-31.
RX MEDLINE=9076671; PubMed=3539202;
RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaut M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol:
conservation across species and homology to platelet IiB/IiIIa.";
RL Biochim. Biophys. Acta 874:369-371(1986).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX MEDLINE=95171458; PubMed=7867070;
RA Lee J.O., Rieu P., Arnaut M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
CR3 (CD11b/CD18).";
RL Cell 80:631-638(1995).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX MEDLINE=96363671; PubMed=8747460;
RA Lee J.O., Bankston L.A., Arnaut M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (I-domain): a pathway for
activation?";
RL Structure 3:1333-1340(1995).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX MEDLINE=98362595; PubMed=9687375;
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrikson R.L.,
Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
Mutchler V.T., Tonich C.S., Watenpaugh K.D., Wiley V.H.;
RT "Cation binding to the integrin CD11b I domain and activation model
assessment.";
RL Structure 6:923-935(1998).
RN [12]
RP 3D-STRUCTURE MODELING OF 17-616.
RX MEDLINE=98226734; PubMed=9560195;
RA Oxvig C., Springer T.A.;
RT "Experimental support for a beta-propeller domain in integrin alpha-
subunits and a calcium binding site on its lower surface.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES
AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF
THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
PEPTIDE IN C3b. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES
OF FIBRINOGEN GAMMA CHAIN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M
ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD11b entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11b.htm".
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to licenses@isb-sib.ch).
CC -----
DR EMBL; J03925; AAA59544.1; -
DR EMBL; M18044; AAA59491.1; -
DR EMBL; J04145; AAA59903.1; -
DR EMBL; S52227; AAB24821.1; -
DR EMBL; S52152; AAB24821.1; JOINED.
DR EMBL; S52153; AAB24821.1; JOINED.
DR EMBL; S52154; AAB24821.1; JOINED.
DR EMBL; S52155; AAB24821.1; JOINED.
DR EMBL; S52157; AAB24821.1; JOINED.
DR EMBL; S52159; AAB24821.1; JOINED.
DR EMBL; S52161; AAB24821.1; JOINED.
DR EMBL; S52164; AAB24821.1; JOINED.
DR EMBL; S52165; AAB24821.1; JOINED.
DR EMBL; S52167; AAB24821.1; JOINED.
DR EMBL; S52169; AAB24821.1; JOINED.
DR EMBL; S52170; AAB24821.1; JOINED.
DR EMBL; S52173; AAB24821.1; JOINED.
DR EMBL; S52174; AAB24821.1; JOINED.
DR EMBL; S52180; AAB24821.1; JOINED.
DR EMBL; S52181; AAB24821.1; JOINED.
DR EMBL; S52184; AAB24821.1; JOINED.
DR EMBL; S52189; AAB24821.1; JOINED.
DR EMBL; S52191; AAB24821.1; JOINED.
DR EMBL; S52192; AAB24821.1; JOINED.
DR EMBL; S52203; AAB24821.1; JOINED.
DR EMBL; S52212; AAB24821.1; JOINED.
DR EMBL; S52213; AAB24821.1; JOINED.
DR EMBL; S52216; AAB24821.1; JOINED.
DR EMBL; S52219; AAB24821.1; JOINED.
DR EMBL; S52220; AAB24821.1; JOINED.
DR EMBL; S52221; AAB24821.1; JOINED.
DR EMBL; S52222; AAB24821.1; JOINED.
DR EMBL; S52226; AAB24821.1; JOINED.
DR EMBL; M76724; AAA58410.1; -
DR EMBL; M84477; AAA51960.1; -
DR PIR; A31108; RWHUIB.
DR PDB; 1A8X; 17-JUN-98.
DR PDB; 1BHO; 18-NOV-98.
DR PDB; 1BHQ; 18-NOV-98.
DR PDB; 1IDN; 25-NOV-98.
DR PDB; 1IDO; 01-AUG-96.
DR PDB; 1JLM; 11-JAN-97.
DR PDB; 1MIU; 07-AUG-02.
DR Genew; HGNC:6149; ITGAM.
DR MM; 120980; -
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion; TAS.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3_
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWFA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Signal; 3d-structure; Repeat; Magnesium; Calcium.
FT SIGNAL 1 16
FT CHAIN 17 1152 INTEGRIN ALPHA-M.
FT DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1105 1128 POTENTIAL.
FT DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).

Query Match		99.6%; Score 5852.5; DB 1; Length 1152;	
Best Local Similarity		99.6%; Pred. No. 0;	
Matches 1132; Conservative		1; Mismatches	1; Gaps
Qy	1	FNLDTENAMTQENARGCQSVQVQOGSRVVVGAPQEIIVAAVNRQSLYQCDYSTGSCPEI	60
Db	17	FNLDTENAMTQENARGCQSVQVQOGSRVVVGAPQEIIVAAVNRQSLYQCDYSTGSCPEI	76
Qy	61	RLQVPEAVNMSLGLSLAATSPQALLACGPTVHQTCSNTYVVKGLCFPLGNSLRQPOK	120
Db	77	RLQVPEAVNMSLGLSLAATSPQALLACGPTVHQTCSNTYVVKGLCFPLGNSLRQPOK	136
Qy	131	FPFALRGCPQEDSDIAFLIDGSGSIIPHDPRMKGFTVSTWBLKXSLTSLWQYSEEP	180
Db	137	FPFALRGCPQEDSDIAFLIDGSGSIIPHDPRMKGFTVSTWBLKXSLTSLWQYSEEP	196
Qy	181	RHPTFKFQNNPNRSLVKPTIQLLGRTHATGVRKIVIRIELNITNGARKNAFKILVVI	240
Db	197	RHPTFKFQNNPNRSLVKPTIQLLGRTHATGVRKIVIRIELNITNGARKNAFKILVVI	256
Qy	241	TGGEKFGDPLGYEDVPEADRGVIRYVIGGDAPRSEKSRQELNTIASKPRDRHVQVN	300
Db	257	TGGEKFGDPLGYEDVPEADRGVIRYVIGGDAPRSEKSRQELNTIASKPRDRHVQVN	316
Qy	301	NPEALKTIONQLREXTAETBGTCTGSSSPHEMSQEGPSAATGNGPLLSVWGSYDWAG	360
Db	317	NPEALKTIONQLREXTAETBGTCTGSSSPHEMSQEGPSAATGNGPLLSVWGSYDWAG	376
Qy	361	GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVIGAPYOHIGLVAMFR	420
Db	377	GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVIGAPYOHIGLVAMFR	436
Qy	421	QNTGWESNANUKGTGIGAFGASLCSVDVDSNGSDTLVGLICAPHYETRGQVSVCP	480
Db	437	QNTGWESNANUKGTGIGAFGASLCSVDVDSNGSDTLVGLICAPHYETRGQVSVCP	496
Qy	481	PRGORARWQCDVLYGEQCPQGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF	540
Db	497	PRG-RARWQCDVLYGEQCPQGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF	555
Qy	541	HGTSGSGISPSHSORAGSKLSPLOYFGCSLSGGQDLTMDGLVDLTVCQAQGHVLLRSQ	600
Db	556	HGTSGSGISPSHSORAGSKLSPLOYFGCSLSGGQDLTMDGLVDLTVCQAQGHVLLRSQ	615
Qy	601	PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLVHVKSTRDLRLEGQIQSVVT	660
Db	616	PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLVHVKSTRDLRLEGQIQSVVT	675
Qy	661	YDLALDSGRPHSAVFNETHKNTROTQVJGLTQTCETLKLQPLNCIEDPVPSPVLRNLF	720
Db	676	YDLALDSGRPHSAVFNETHKNTROTQVJGLTQTCETLKLQPLNCIEDPVPSPVLRNLF	735
Qy	721	SLVGTPLSAGNLRPVLAEADQRLFTALPFPEKNCNDNIQDDLSITPFSMSLDCLVVG	780
Db	736	SLVGTPLSAGNLRPVLAEADQRLFTALPFPEKNCNDNIQDDLSITPFSMSLDCLVVG	795
Qy	781	GPREFNVTVVRNDGEDSYRTQVTPFPPLDLSYRKVSTLQNSQSRSLACESASSTEV	840
Db	796	GPREFNVTVVRNDGEDSYRTQVTPFPPLDLSYRKVSTLQNSQSRSLACESASSTEV	855
Qy	841	SGALKSTSCSINHPIEPENSEVFNITFDVDSKASLGNKLLKANTYSENMPRTNKTIF	900
Db	856	SGALKSTSCSINHPIEPENSEVFNITFDVDSKASLGNKLLKANTYSENMPRTNKTIF	915
Qy	901	QLELPVKYAYVMVTHSGVSTKYLNFNTASNTSRVMQHQYQVSNLQORSPLSLVLPV	960
Db	916	QLELPVKYAYVMVTHSGVSTKYLNFNTASNTSRVMQHQYQVSNLQORSPLSLVLPV	975
Qy	961	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAEKAPVNCSTAVCORICDIP	1020
Db	976	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAEKAPVNCSTAVCORICDIP	1035

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X07640; CAA30479.1; -;
 DR EMBL; ML4293; AAA39484.1; -;
 DR PIR; S00551; S00551.
 DR HSSP; P11215; 1A8X.
 DR MGD; MGI:96607; Itgam.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PRO1185; INTEGRINA.
 DR PRINTS; PRO0453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS00234; VWA; 1.
 DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Calcium; Repeat.
 FT SIGNAL 1 16
 FT CHAIN 17 1153 INTEGRIN ALPHA-M.
 FT DOMAIN 17 1105 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1106 1123 POTENTIAL.
 FT DOMAIN 1130 1153 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 31 84 FG-GAP 1.
 FT REPEAT 7 7 FG-GAP 2.
 FT DOMAIN 164 350 VWA.
 FT REPEAT 337 400 FG-GAP 3.
 FT REPEAT 401 452 FG-GAP 4.
 FT REPEAT 454 515 FG-GAP 5.
 FT REPEAT 517 575 FG-GAP 6.
 FT REPEAT 580 632 FG-GAP 7.
 FT CA_BIND 465 473 POTENTIAL.
 FT CA_BIND 529 537 POTENTIAL.
 FT CA_BIND 592 600 POTENTIAL.
 FT SITE 1132 1136 GPFKR MOTIF.
 FT DISULFID 66 73 BY SIMILARITY.
 FT DISULFID 105 123 BY SIMILARITY.
 FT DISULFID 654 711 BY SIMILARITY.
 FT DISULFID 770 776 BY SIMILARITY.
 FT DISULFID 999 1023 BY SIMILARITY.
 FT DISULFID 1028 1033 BY SIMILARITY.
 FT CARBOHYD 58 59 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 391 391 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 696 696 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 734 734 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 772 772 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 801 801 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 881 881 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 907 907 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 941 941 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 980 980 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 994 994 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1022 1022 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1051 1051 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1076 1076 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1153 AA; 127480 MW; 178DB988AECB0343 CRC64;

Query Match 76.2%; Score 4476; DB 1; Length 1153;
 Best Local Similarity 74.3%; Pred. No. 2.6e-290;
 Matches 845; Conservative 142; Mismatches 149; Indels 2; Gaps 2;
 1 FNLDTENAMTFQENARGFGQVYVQLOGSRVVGAPQEIIVAAHQGSLYQCYDSTGSCBPI 60
 17 FNLDTEHPMTQENAKGFGQVYVQLOGSRVVGAPQEIIVAAHQGSLYQCYDSTGSCBPI 76

QY 61 RLQVPEAVNMVSLGLSLAATTPPQLLACGPTVHQTSENVTYKGLCLFPGSNLRQPOK 120
 DB 77 PLQVPEAVNMVSLGLSLAVSTVPQQLLACGPTVHQTSENVTYKGLCLFPGSNLRQPOK 136
 QY 121 PPEALRGCPQEDSDIATLIDGSGSIIPHDPRMEKEFVSTWEOQLFKSKTLFSLMQYSEF 180
 DB 137 PPEALRGCPQEDSDIATLIDGSGSIIPHDPRMEKEFVSTWEOQLFKSKTLFSLMQYSEF 196
 QY 181 RIHFTFXEQNNPNRSLVKEITQLLGRTHATATVRKVIKRELLNITNGARKNAKILVI 240
 DB 197 RIHFTFXEQNNPNRSLVKEITQLLGRTHATATVRKVIKRELLNITNGARKNAKILVI 256
 QY 241 TDGKFGDPLGYEDVIPEADREGVIRVYVGVCDAPRSEKSHOELATITASKPRDHVPOVN 300
 DB 257 TDGKFGDPLGYEDVIPEADREGVIRVYVGVCDAPRSEKSHOELATITASKPRDHVPOVN 316
 QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSEFHEMSQGFSAITENGPLLSTVGSYDWAG 360
 DB 317 NFEALKTIQNLREKIFAIEGTQTGSSSEFHEMSQGFSAITENGPLLSTVGSYDWAG 376
 QY 361 GVFLYTSKEKSTFIMTRVDSMDADAYLGAAAIILNRVQSLVIGAPRYOHIGLVAMFR 420
 DB 377 GAFLYTSKDVTFINTTRVDSMDADAYLGAAAIILNRVQSLVIGAPRYOHIGLVAMFR 436
 QY 421 QNTGWSNANVKTQIGAVFGASLCSVDVDSNGSTDLVIGAPRYOETQTRGGQVSVCP 480
 DB 437 ENFGTWEPHTSIKGSQISYFGASLCSVDVDSNGSTDLVIGAPRYOETQTRGGQVSVCP 496
 QY 481 PRGQARWQCDVILYGGQGPWGFAGALTVLGVNKGKLTDAVIGAPGEDNRGAVLP 540
 DB 497 PRG-RARWQCEALLHGDQHPWGFAGALTVLGVNKGKLTDAVIGAPGEDNRGAVLP 555
 QY 541 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLGQDLTMDGLVDLTGACGHVLLRSQ 600
 DB 556 YGASTASLASHSHRIIGAHFSPGLQYFGQSLGQDLTMDGLVDLTGACGHVLLRSQ 615
 QY 601 PVLRYKAIMFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDRLEGOIQSVVT 660
 DB 616 PVLRYKAIMFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDRLEGOIQSVVT 675
 QY 661 YDLALDSGRPHSRAVENETKSTRTOVGLTCTETKLQLPNCITBDPVSIVLRNF 720
 DB 676 YDLALDSGRPHSRAVENETKSTRTOVGLTCTETKLQLPNCITBDPVSIVLRNF 735
 QY 721 SLVGTPLSAFGLRPLVLAEDAQRLLFTALFPFERKCGNDNICODDI:SIITFSPLMCLVVG 780
 DB 736 TLVGEPLRSFGLRPLVLAEDAQRLLFTALFPFERKCGNDNICODDI:SIITFSPLMCLVVG 795
 QY 781 GPRFPNVTVRNDEGDSYRTOVTFFPPLDLSYKXVSTLQNRSCRSWRL-ACBSASSTE 839
 DB 796 GPRFPNVTVRNDEGDSYRTOVTFFPPLDLSYKXVSTLQNRSCRSWRL-ACBSASSTE 855
 QY 840 VSGALKSTSCSINHPPIPPENSEVTNITFDVDSKASLGNKLLK:NVTSNNMPTNKTE 899
 DB 856 VSGALKSTSCSINHPPIPPENSEVTNITFDVDSKASLGNKLLK:NVTSNNMPTNKTE 915
 QY 900 FQLELPKYAVYVWVTSHGVTKYLNFTASENTGRVMQHQVQVNLGORSIPISVFLVP 959
 DB 916 FQLELPKYAVYVWVTSHGVTKYLNFTASENTGRVMQHQVQVNLGORSIPISVFLVP 975
 QY 960 VRLNQTVLWDRPQVTFSENLSSTCTHKEKRLPSHSDFLAELKAPVWNCISVACORICDI 1019
 DB 976 VRLNQTVLWDRPQVTFSENLSSTCTHKEKRLPSHSDFLAELKAPVWNCISVACORICDI 1035
 QY 1020 PFFGQESFNATLKGNSFDWYIKTSHNHLIIVSTAEILLFNDVSFTLLPGQAGFVSOTE 1079
 DB 1036 PFFGQESFNATLKGNSFDWYIKTSHNHLIIVSTAEILLFNDVSFTLLPGQAGFVSOTE 1095
 QY 1080 TKVPEFYPNPLKLVGSGVGLLALLIITAAVLKGFPEKQYKMMSEGGPPGAPQ 1137
 DB 1096 TKVPEFYPNPLKLVGSGVGLLALLIITAAVLKGFPEKQYKMMSEGGPPGAPQ 1153

RESULT 3
 ITAX_HUMAN STANDARD; PRT; 1163 AA.
 AC P20702;
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
 DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Jeu M5).
 GN ITGAX OR CD11C
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8816645; PubMed=3327687;
 RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
 RT "cDNA cloning and complete primary structure of the alpha subunit of
 RT a leukocyte adhesion glycoprotein, p150,95."
 RL EMBL J. 6:4023-4028 (1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90153906; PubMed=2303426;
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
 RL J. Biol. Chem. 265:12750-12751 (1990).
 RN [4]
 RP SEQUENCE OF 20-43.
 RX MEDLINE=87167596; PubMed=3549901;
 RA Miller L.J., Wiebe M., Springer T.A.;
 RT "Purification and alpha subunit N-terminal sequences of human Mac-1
 RT and p150,95 leukocyte adhesion proteins."
 RL J. Immunol. 138:2381-2383 (1987).
 CC -!- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT
 CC RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL
 CC INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY
 CC IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X
 CC ASSOCIATES WITH BETA-2.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
 CC GRANULOCYTES.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide Cdlc entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cdlc.htm".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M81695; AA55180.1; -;
 CC EMBL; Y00093; CA68283.1; -;
 CC EMBL; M29185; -; NOT ANNOTATED CDS.
 CC EMBL; M29487; AA51620.1; ALT SEQ.
 CC EMBL; M29482; AA51620.1; JOINED.
 CC EMBL; M29483; AA51620.1; JOINED.
 CC EMBL; M29484; AA51620.1; JOINED.
 CC EMBL; M29485; AA51620.1; JOINED.

EMBL; M29486; AA51620.1; JOINED.
 PIR; A36584; RWHUIC.
 PDB; 1N3Y; 18-FEB-03.
 Genew; HGNC:6152; ITGAX.
 DR MIM; 151510; -;
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 DR GO; GO:0004872; F:receptor activity; TAS.
 DR GO; GO:0007155; P:cell adhesion; TAS.
 DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR020335; VWFA.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_A; 1.
 DR Pfam; PF00692; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS0234; VWFA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Magnesium; Calcium; Repeat; 3D-structure.
 FT SIGNAL 1 19 INTEGRIN ALPHA-X.
 FT CHAIN 20 1163 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 20 1107 POTENTIAL.
 FT TRANSMEM 1108 1128 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 1129 1163 FG-GAP 1.
 FT REPEAT 34 87 FG-GAP 2.
 FT REPEAT ? ? VWFA.
 FT DOMAIN 165 351 FG-GAP 3.
 FT REPEAT ? ? FG-GAP 4.
 FT REPEAT 402 453 FG-GAP 5.
 FT REPEAT 455 517 FG-GAP 6.
 FT REPEAT 518 576 FG-GAP 7.
 FT REPEAT 581 633 POTENTIAL.
 FT CA_BIND 466 474 POTENTIAL.
 FT CA_BIND 530 538 POTENTIAL.
 FT CA_BIND 593 601 POTENTIAL.
 FT SITE 1131 1135 GFFKR MOTIF.
 FT DISULFID 69 76 BY SIMILARITY.
 FT DISULFID 108 126 BY SIMILARITY.
 FT DISULFID 655 712 BY SIMILARITY.
 FT DISULFID 771 777 BY SIMILARITY.
 FT DISULFID 848 863 BY SIMILARITY.
 FT DISULFID 998 1022 BY SIMILARITY.
 FT DISULFID 1027 1032 BY SIMILARITY.
 FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 697 697 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 735 735 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 899 899 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 939 939 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1050 1050 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 490 490 G -> A (IN REF. 2).
 FT CONFLICT 756 756 L -> D (IN REF. 2).
 SQ SEQUENCE 1163 AA; 127885 MW; 6C4E19CC3P62A473 CRC64;
 Query Match 59.1%; Score 3473; DB 1; Length 1163;
 Best Local Similarity 61.3%; Pred. No. 2,1e-223;
 Matches 692; Conservative 138; Mismatches 293; Indels 6; Gaps 4;
 Qy 1 FNLDTEVMTQENARFGSGSVVQLQSGSVVVGAPQPIVAVANQRCGLYQCDYSTGSCPEI 60
 Db 20 FNLDTEBELTAPRVDSAGFSDSVVQVANSVWVVGAPQKITAANQTGSGYQCGYSTGACEPI 79
 Qy 61 RLQVPVEAVNMSLGLSLAAATTPPQLLACGPTVHQTCSNTYVKGCLFLFSGNLRQOPQK 120
 Db 80 GLQVPPEAVNMSLGLSLASTSPQLLACGPTVHHCGRNWLVLGCLLIGPT--QLTOR 137
 Qy 121 FPEALRCQPQEDSDIAFLIDQSGSIIPHDFRMKEFVSTVMEQLKSKTLPSLQYQSEFP 180

Df 138 LPVSRQECPRQBDIVFLIDSGSISRRNFATWVRAVISQFQRPBPSTQFSLMQFSNKP 197
Qy 181 RHFTFKERFONNPNPESLVKPTOLLGRTHRTATGVRKVIPELLNITNGARKNAKILLIVI 240
Df 198 QHFTFEPRFRTNPLSLASVHQLQGFYTTATLQNVVHRFLFHASYGARDATKILLIVI 257
Qy 241 TDCEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPDRHVPQV 300
Df 258 TDGKEGSDLVKDVIPMDADAIIRYAIGVGLAFQNRNSWKEINDIASKPSQEHFKVE 317
Qy 301 NFEALKTIONLREKIPATECTQTSSEFHEMSEGEFSNAITNGSLTSSVGYDAG 360
Df 318 DFDALKDIONLREKIPATECTQTSSEFHEMSEGEFSNAITNGSLTSSVGYDAG 377
Qy 361 GYFLVTSKESKTFINWTRVDSMDAYLGYAAAILRNVRQSLVGLGAPRYQHIGLVAFPR 420
Df 378 GAFLYPPNMSPTFINMSQENVDNRSDYLCYSTELALWKGVSQSLVGLGAPRYQHTGRAVFT 437
Qy 421 QNTGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVILICAPHYVETRCQGVSVCP 480
Df 438 QVSRQWRMKAETVGTQIGSYFGASLCSVDVDTGSDTDLVILICAPHYVETRCQGVSVCP 497
Qy 481 PRGQARWOCDAVLVGEQCPWRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVLYF 540
Df 498 PRGWR-RWMCDAVLVGEQCPWRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVLYF 556
Qy 541 HGTSGGSISSHSQRIAGSKLSRLOYFCQSLGGQDLNMGDLNLTGCAQHVILLRSQ 600
Df 557 HGVLGPSISFSHSQRIAGSKLSRLOYFCQSLGGQDLNMGDLNLTGCAQHVILLRSQ 616
Qy 601 PVLKVKAIEMPREVARNVFECNDQVKGKEAGEVVRVCLHVOKSTRDLRGQIQSVVT 660
Df 617 PVLWGVSVHQFTPAETPRSAFECREOVSEQTLVQSNICLYIDKSKILLGSRDLQSVT 676
Qy 661 YDLALDSGRPHRAVNETKSTRQTVLGTQTCETKLPNCIEDPVPVILRLNF 720
Df 677 LDALDPGLSPRATFOETKNSLSRVRVLGLKAHCENFNLLPSCEDSVTPITRLNF 736
Qy 721 SVGTLSAPGNLSPVLAEDALEFTALPFPKNGCNDNICODDLSITSEKSLDCLVVG 780
Df 737 TLVGKPELLAFRLNRLPMLAALORYFTASUPFEKNCAGDHCQDNLGSISSFGLKSLVVG 796
Qy 781 GPREFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRORSQMSKLACESASTEV 840
Df 797 SNLELNAEVMWVNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRORSQMSKLACESASTEV 854
Qy 841 SGALKSTCSINHPISPESEVTFNITFDVDSKASLGKLLKLLKANVTSENMPRTKTF 900
Df 855 SQGTWSTSCRIHLPFGGAQITFLATFDVSPKAVLGDRLLLTANVSSENTPRTKTF 914
Qy 901 QLELPVKYAVMYVTVSHGVSTKYLNFAS-ENTSRVWQHQYQVSNLQGRSLPISLFLVP 959
Df 915 QLELPVKYAVTVVSSHQFTKYLNFSESEKESHVAMHYQVNNLQGRDLPVSNFWVP 974
Qy 960 VRLNQVIMDRPOTVSENLSSTCHTKERLPSHDFLAELKAPVNVCSIAVCORIQDI 1019
Df 975 VELNQAVMNDVEVSHFPQNPDLSCSEKIAFPASDFLAHQNPVLDSCIAGLRCFCDV 1034
Qy 1020 PFGIOEFNATLKNLSPDWYIKTSHNLLIYSTABILFNDVSFTLLPQCAQVRSQTE 1079
Df 1035 PFSVQELDFTLKNLSPDWYIKTSHNLLIYSTABILFNDVSFTLLPQCAQVRSQTE 1094
Qy 1080 TKVEPVPNPLIVGVSSVGGILLALLALITAAVLYKLGFFKRYQKWMSE 1128
Df 1095 TULEKVKVHNFTLIVGVSSVGGILLALLALITAAVLYKLGFFKRYQKWMSE 1143

RESULT 4
ITAD HUMAN STANDARD; PRT; 1162 AA.
ID ITAD HUMAN Q13349; Q15576;
AC Q13349; Q15576;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)
Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
ITGAD
Homo sapiens (Human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=Spleen;
MEDLINE=96111956; PubMed=8777714;
Van der Vliet M., Le Trong H., Wood C.L., Moore P.F., St John T.,
Stanton D.E., Gallatin W.M.;
"A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-
3";
J. Biol. Chem. 275:8959-8969(2000).
[3]
SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
MEDLINE=96257236; PubMed=8666289;
Wong D.A., Davis E.M., Lebeau M., Springer T.A.;
"Cloning and chromosomal localization of a novel gene encoding a human
beta 2-integrin alpha subunit";
Gene 171:291-294(1996).
[4]
INTERACTION WITH VCAM1.
MEDLINE=95059842; PubMed=9841932;
Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W.,
Hoffman P.A., Stanton D.E., Bochner B.S.;
"alpha2 integrin is expressed on human eosinophils and functions
as an alternative ligand for vascular cell adhesion molecule 1
(VCAM-1)";
J. Exp. Med. 188:2187-2191(1998).
[5]
INTERACTION WITH VCAM1.
MEDLINE=95370002; PubMed=10438935;
Van der Vliet M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,
Grayson M.H., Bochner B.S., Gallatin W.M., Stanton D.E.;
"The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a
binding interface between I domain and VCAM-1";
J. Immunol. 163:1984-1990(1999).
- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND
VCAM1. MAY PLAY A ROLE IN THE ATHEROSCLEROTIC PROCESS SUCH AS
CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-
BORNE PATHOGENS, PARTICULATE MATTER, AND SENESCENT ERYTHROCYTES
FROM THE BLOOD.
- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D
ASSOCIATES WITH BETA-2.
- SUBCELLULAR LOCATION: Type I membrane protein.
- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL
LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON
TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FCAM CELLS WITHIN
ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.
- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
- SIMILARITY: Belongs to the integrin alpha chain family.
- SIMILARITY: Contains 1 VWFA domain.
- SIMILARITY: Contains 7 FG-GAP repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
or send an email to license@sib-sib.ch.

DR	EMBL; U37028; AAB38547.1; -	Qy	1	FNLDENAMTFOENARQFGOSVVOLOQSRVWVGAPQBIIVAAQBSGLYQCDYSTGSCPEI	60
DR	EMBL; U40274; AAB60634.1; -	Db	18	FNLDVEEFTTFOEDAGGFGOSVVOFGSRLVWVGAPLEVAANQRTDYDCAAAATGCOPI	77
DR	EMBL; U40275; AAB60635.1; -	Qy	61	RLOVPEAVANMSLGLSLAATTSPPQIACGPTVHOTCSENTYVKLCFLFGSNLRQOPQK	120
DR	EMBL; U40276; AAB60636.1; -	Db	78	PLHIRPEAVANMSLGLTLAASNGSRLLACGPTLHRVCGENSYSKSGCLLLGSRW-ELIQT	136
DR	EMBL; U40277; AAB60637.1; -	Qy	121	FPALRCPOEDSDIAPLIDGSGSIIPHDFRMKEFVSTWMEQLKSKTFLPSLMQYSEEP	180
DR	EMBL; U40278; AAB60638.1; -	Db	137	VPDAPTECPHQEMDIVFLIDGSGSIDQDNFNQKGFQAVMGQEGFTDTTLPALQOYNNLL	196
DR	HSSP; P11215; IABX.	Qy	181	RIHPTREFOFNNPRSLVKPIITOLLGRTHATGVRKRVIRELLNITNGARKNAFKILIVI	240
DR	MIM; 602453; -	Db	197	KIHFTFTQFTSPSQSLVDPIVOLKLTFTATGILTVVTLQFHHKNGARKSACKILIVI	256
DR	GO; GO:0008305; C:integrin complex; TAS.	Qy	241	TDGEKFGDPLGYEDVIPEDAREGVIRYVIGVDAFRSEKSEKQELNITASPPRHHVQVN	300
DR	GO; GO:0016337; P:cell-cell adhesion; NAS.	Db	257	TDGQYKDPLEYSQVIPAOKAGIIRYVIGVDAFRSEKSEKQELNITASPPRHHVQVN	316
DR	GO; GO:0006955; P:immune response; NAS.	Qy	301	NFEALKTIONQLEKIPALIGTGTGSSSSFEHMSQSGFSAATITSNPLSTVGSYDWAG	360
DR	InterPro; IPR000413; Integrin_alpha.	Db	317	NFAALGSIQKLOQEKIYAVEGTQSRASSSQHMSQSGFSTALTMDGLFLGAVGSPFWSG	376
DR	InterPro; IPR002035; VWF_A.	Qy	361	GVFLYTSKEXSTFINMTRVDSMDNDYGLYAAAIILENNRVQSLVGLGAPRYQHIGLVAMFR	420
DR	Pfam; PF00357; integrin_A; 1.	Db	377	GAPLYPPNMSPTFINMSQENVDMEDSYLGYSTELALWQONLVGLGAPRYOHTGKAVFT	436
DR	Pfam; PF00092; vwa; 1.	Qy	421	QNTGHWESNANVKTQIGAVFGASLCSDVDSNGSTDLVLIGAVYVEOTEQGGVSVCP	480
DR	PRINTS; PRO1185; INTEGRINA.	Db	437	QVSRQWKAQEVGTQIGSYFGASLCSDVDSNGSTDLVLIGAVYVEOTEQGGVSVCP	496
DR	SMART; SM00191; Int alpha; 4.	Qy	481	PRGORAEWQCDVLYGEGQGPWGRFGAALTVLGVNGDKLTDVAIGAPGEDNKGAVLYF	540
DR	SMART; SM00327; VWA; 1.	Db	497	PRGORVQWQCDVLYGEGQGPWGRFGAALTVLGVNGDKLTDVAIGAPGEDNKGAVLYF	556
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.	Qy	541	HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLSGQDLTMDGLVDLTVGAOCHVLLRSQ	600
DR	Signal; Repeat; Calcium;	Db	557	HGASESGISPSHSQRIASSQLSPLOYFGQALSGQDLTMDGLVDLTVGAOCHVLLRSQ	616
DR	Magnesium.	Qy	601	PVLAKVKAIMEFNPREVARNVPECDNDVVVKGKAGEVRVCLHVQKSTRDLAEQIGTQSVV	660
DR	CHAIN	Db	617	PVLKVGVMRFPVPEVAKAVYRCWEKESALEAGDATTCLTIQKSLDQL--GDIQSSVR	674
DR	DOMAIN	Qy	661	YDLALDSGRPHSRVAFNETKSTRTQVIGLGTCTETLKLQPNCHIDPVPVILNLF	720
DR	REPEAT	Db	675	FDLALDQRLTSRAIFNETKNPTLTRKTLGLGHCHETLKLLEPCVEDVVSPIILHNF	734
DR	REPEAT	Qy	721	SLVGTPLSAFGLNRPVLAEDAQRLFTALFPKKNCGNDNICODDLISITFSFMSLDCLVVG	780
DR	REPEAT	Db	735	SLVREPIPSQNLPRPVLAVGSDQLFTASLPFKKCGQDGLCEGBLVTLSFSGLTTLVG	794
DR	REPEAT	Qy	781	GPREFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASTEV	840
DR	REPEAT	Db	795	SSLELNVLTVVWVNGEDSYGTWVSLYPAGLSHRVSGAQKQPHQSALRLACETV-PTED	853
DR	REPEAT	Qy	841	SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMMPTNKTEF	900
DR	REPEAT	Db	854	EG-LRSSRCVSNHPIFHHGSGNGTFTVTFDVSFKATLGRDMLMRFAESSENKASSKATP	912
DR	REPEAT	Qy	901	QLELPVKYAVTMVWVTSHGVSSTKYNLF-TASENTSRVMOHYQVNSLQORSPLISLVFLVP	959
DR	REPEAT	Db	913	QLELPVKYAVTMVWVTSHGVSSTKYNLF-TASENTSRVMOHYQVNSLQORSPLISLVFLVP	972
DR	REPEAT	Qy	960	VLNQTQVTDPRQVTFPSNLSSTCHTKERLPSHSDFLAELKAPVNVCSIAVCQRIODI	1019
DR	REPEAT	Db	973	VLLNGVAVDVVMEAPSQSL--PCVSEKRPQHSDFLTQISRSPLWLDLCSIDACQIFRCDV	1030
DR	REPEAT	Qy	1020	PPFGIOEBFNATLKNLSFDWVTKSHNHLITVSTABILFNDSVFTLLPGQCAFVRSQTE	1079
DR	REPEAT	Db	1031	PSFSVQEBELDTLKNLSFGVNRVETLQCKLVVSVABITFTDTSVYSLQPGQAFPAQWE	1090
DR	REPEAT	Qy	1080	TKVEPFEPNPLPLIVGSSVGLLLALLITLALYKLGFFKQYKQKMSB 1128	

Query Match 58.2%; Score 3417; DB 1; Length 1162;
 Best Local Similarity 59.8%; Pred. No. 1.1e-219;
 Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;

```
DB      1091 MVLBEDEVNAPITMGSSVGAULLLALITATLYKLGFKRHYKEMLED 1139
RESULT 5
IPAL_HUMAN
AC      P20701; O43746;          STANDARD;          PRT; 1170 AA.
DT      01-FEB-1991 (Rel. 17, Created)
DE      01-NOV-1995 (Rel. 32, Last sequence update)
DE      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE      alpha chain) (Leukocyte function associated molecule 1, alpha chain)
DE      (CD11a)
GN      ITGAL OR CD11A.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
RX      MEDLINE=89139587; PubMed=2537322;
RA      Larson R.S., Corbi A.D., Berman L., Springer T.;
RA      "Primary structure of the leukocyte function-associated molecule-1
RT      alpha subunit: an integrin with an embedded domain defining a protein
RT      superfamily.";
RL      J. Cell Biol. 108:703-712(1989).
RN      [2]
RP      SEQUENCE FROM N.A. (ISOFORM 2).
RX      MEDLINE=99425270; PubMed=10493829;
RA      Loftus B.J., Kim U.-J., Sneddon V.P., Kalush P., Brandon R.,
RA      Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
RA      Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
RA      Eichler B.E., Harris P.C., Venter J.C., Adams M.D.;
RT      "Genome duplications and other features in 12 Mb of DNA sequence from
RT      human chromosome 16p and 16q.";
RL      Genomics 60:295-308(1999).
RN      [3]
RP      X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.
RX      MEDLINE=96036067; PubMed=7479767;
RA      Ou A., Leahy D.J.;
RT      "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha
RT      L beta 2) integrin.";
RL      Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281(1995).
RN      [4]
RP      X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.
RX      MEDLINE=96398682; PubMed=8805579;
RA      Ou A., Leahy D.J.;
RT      "The role of the divalent cation in the structure of the I domain
RT      from the CD11a/CD18 integrin.";
RL      Structure 4:931-942(1996).
RN      [5]
RP      X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.
RX      MEDLINE=99425288; PubMed=10493852;
RA      Kallen J., Welzenbach K., Ramage P., Geyl D., Kriwacki R., Legge G.,
RA      Cottens S., Weitz-Schmidt G., Hommel U.;
RT      "Structural basis for LFA-1 inhibition upon lovastatin binding to the
RT      CD11a I-domain.";
RL      J. Mol. Biol. 292:1-9(1999).
CC      -1- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
CC      ICAM3 AND ICAM4. IT IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA
CC      INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL
CC      MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES
CC      AND MONOCYTES.
CC      -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L
CC      ASSOCIATES WITH BETA-2.
CC      -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC      -1- ALTERNATIVE PRODUCTS:
CC      Event=Alternative splicing; Named isoforms=2;
CC      Name=1;
CC      IsoId=P20701-1; Sequence=Displayed;
CC      Name=2;
CC      IsoId=P20701-2; Sequence=VSP_002738;
```

Note=No experimental confirmation available;

-1- TISSUE SPECIFICITY: LEUKOCYTES

-1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

-1- SIMILARITY: Belongs to the integrin alpha chain family.

-1- SIMILARITY: Contains 1 VWFA domain.

-1- SIMILARITY: Contains 7 FG-GAP repeats.

-1- DATABASE: NAME=PROW; NOTE=CD guide CD11a entry; WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11a.htm".

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Y00796; CAA68747.1; --

EMBL; AC062310; AAC31672.1; --

PIR; S03308; S03308.

PDB; 1LFA; 29-JAN-96.

PDB; 1ZON; 07-DEC-96.

PDB; 1ZOO; 07-DEC-96.

PDB; 1ZOP; 07-DEC-96.

PDB; 1CQP; 07-AUG-00.

PDB; 1DQG; 03-FEB-00.

PDB; 1MJN; 28-JAN-03.

PDB; 1MQ8; 14-JAN-03.

PDB; 1MQ9; 14-JAN-03.

PDB; 1MQA; 14-JAN-03.

Genew; HGNC:6148; ITGAL.

MIM; 153370; --

GO; GO:0008305; C:integrin complex; TAS.

GO; GO:0006928; P:cell motility; TAS.

InterPro; IPR000413; Integrin_alpha.

InterPro; IPR002035; VWF_A.

Pfam; PF01839; FG-GAP; 3.

Pfam; PF00357; Integrin_A; 1.

Pfam; PF00092; vwa; 1.

PRINTS; PR01185; INTEGRINA.

PRINTS; PR00453; VWFADOMAIN.

SMART; SM00191; Int_alpha; 4.

SMART; SM00327; VWF_1.

PROSITE; PS00242; INTEGRIN_ALPHA; 1.

PROSITE; PS00234; VWFA; 1.

Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane; Signal; 3D-structure; Magnesium; Calcium; Repeat; Alternative splicing.

SIGNAL 1 25

CHAIN 26 1170 INTEGRIN ALPHA-L.

DOMAIN 26 1088 EXTRACELLULAR (POTENTIAL).

TRANSMEM 1089 1112 POTENTIAL.

DOMAIN 1113 1170 CYTOPLASMIC (POTENTIAL).

REPEAT 42 91 FG-GAP 1.

REPEAT 92 149 FG-GAP 2.

DOMAIN 170 349 VWFA.

REPEAT ? ? FG-GAP 3.

REPEAT 401 455 FG-GAP 4.

REPEAT 457 516 FG-GAP 5.

REPEAT 518 575 FG-GAP 6.

REPEAT 578 630 FG-GAP 7.

CA_BIND 468 476 POTENTIAL.

CA_BIND 530 538 POTENTIAL.

CA_BIND 590 598 POTENTIAL.

SITE 1115 1119 GFEK MOTIF.

DISULFID 73 80 BY SIMILARITY.

DISULFID 111 129 BY SIMILARITY.

DISULFID 653 707 BY SIMILARITY.

DISULFID 771 777 BY SIMILARITY.

DISULFID 845 861 BY SIMILARITY.

DISULFID 998 1013 BY SIMILARITY.

DISULFID 1021 1052 BY SIMILARITY.

Db 599 VCAEGRVTVLSSRPVVDVVTLSFSPBEPVHEVCYSAREEQHGKVLKACRAIKPLT 658
Qy 647 RDLREGQIQSVTVYDLALDSCRPESRAVENETKNSRTRQVGLGTOTCTELKQLPNC 706
Db 659 PQ--FQGRLLANLSTLQDHRMSRGLFPGDSHELNGTSITP-DKSCLDLFFHFPIC 715
Qy 707 IEDPVSPVILRNFLSV---GTPISAFGN-LRPVLAEDAQRLFTALPFFKNCNDNICQ 762
Db 716 IQDLSPINVLNLSLLEEGTPQKGRAMPILRPSIHTV-TKEIFPKNCGEDKKCE 774
Qy 763 DDLSTPFSMSLDCLVVGSP-----REFNVTVTVRNDGDSYRQVTFPPPLDLVSRKV 816
Db 775 ANLTUSSPARS-----GPLRMSSASLAVETLUSNGEDAYVWLDDLPFRGLSPRKV 827
Qy 817 STLQNRQSRGRLACESASSTEVSGAL-KSTSCSINHPIPPENSEVTFNITFVDVSKAS 875
Db 828 EMLO---PHSRMPVSCBEL--TEGSSLLTKLKNVSPPIKAGQEVSLQVWFNTLLMS 882
Qy 876 LONKLLIXANTISEN-NMPTNKTEFQLELPKVIYVWVTSHGVSITKLNFTASENISR 934
Db 883 WEDFVELNGTVHCENENSSLOEDNSAATHIPVLYPWNILTKQENSTLYISFTPKGPKTQ 942
Qy 935 VQHOYQVSNLQORSLPISLFLVPLVRLNQTIVMDRPO---VTPSENLSS---TCHTK 986
Db 943 QVQHYQV-----RQPSAYDHMET-LEALGVPRPHSEDLITYTNSVQTDPLVTCHE 996
Qy 987 E-RLPSHSDFLAELKAPVNCISAVCORIQCDIPFGIOBEFNATLKNLSFDWYIKTS 1045
Db 997 DLKRPSSB---AEQCLFGV-----QFRCPIVF---RWEIITQVGTVELSKEIKAS 1042
Qy 1046 HNHLLIVSTAELFNDSTVFTLLPGOGAFVRSQTEKVPPEVNPPLPLGVSSVGLLL 1105
Db 1043 -STLSLCSGLSVSNSSKHFLYKSKA-SEAQLVKVLDLHHEKMLHYVVLVSGIGLVLL 1100
Qy 1106 ALITAAKYLGFFKQYKDMM-SEGPPGAEP 1136
Db 1101 FLIFLALYKVGFKXNLKEMKADGVNGSP 1132
RESULT 7
ITAE MOUSE STANDARD; PRF; 1167 AA.
AC Q60677;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Integrin alpha M290).
GN ITGA3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=AKR;
RX MEDLINE=95187992; PubMed=7882170;
RA Smith T.J., Ducharme L.A., Shaw S.K., Parker C.M., Brenner M.B.,
RA Kilshaw P.J., Weis J.H.;
RT "Murine M290 integrin expression modulated by mast cell activation.";
RL Immunity 1:393-403(1994).
CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
CC CELL MONOLAYERS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-E
CC SUBUNIT GENE EXHIBIT A MARKED REDUCTION IN THE NUMBERS OF
CC INTRAEPITHELIAL LYMPHOCYTES IN THE GUT AND IN THE DEVELOPMENT OF
CC GUT-ASSOCIATED LYMPHOID AGGREGATES, SUPPORTING A SPECIFIC ROLE FOR
CC THIS INTEGRIN IN MEDIATING RETENTION OF LYMPHOCYTES IN THE
CC INTESTINAL WALL.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U12236; AAC52142.1; -;
CC HSP; P11215; IABX.
CC MGD; MGI:1298377; Itgae.
CC InterPro; IPR00413; Integrin_alpha.
CC InterPro; IPR002035; VWF_A.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; Integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC PRINTS; PR01185; INTEGRINA.
CC PRINTS; PR00453; VWFADOMAIN.
CC SMART; SMO0191; Int_alpha; 3.
CC SMART; SMO0327; VWA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS00234; VWFA; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Signal; Repeat; Magnesium;
CC Calcium.
CC SIGNAL 1 19
CC CHAIN 20 1167 BY SIMILARITY.
CC CHAIN 20 181 INTEGRIN ALPHA-E.
CC CHAIN 183 1167 INTEGRIN ALPHA-E LIGHT CHAIN.
CC DOMAIN 20 1114 INTEGRIN ALPHA-E HEAVY CHAIN.
CC TRANSMEM 1115 1137 EXTRACELLULAR (POTENTIAL).
CC DOMAIN 1138 1167 POTENTIAL.
CC REPEAT ? ? CYTOPLASMIC (POTENTIAL).
CC REPEAT ? ? FG-GAP 1.
CC REPEAT ? ? FG-GAP 2.
CC DOMAIN 149 192 X-DOMAIN (EXTRA DOMAIN).
CC DOMAIN 193 384 VWFA.
CC REPEAT ? ? FG-GAP 3.
CC REPEAT 449 501 FG-GAP 4.
CC REPEAT 503 564 FG-GAP 5.
CC REPEAT 566 631 FG-GAP 6.
CC REPEAT 634 686 FG-GAP 7.
CC CA_BIND 514 522 POTENTIAL.
CC CA_BIND 578 586 POTENTIAL.
CC DOMAIN 646 654 POTENTIAL.
CC SITE 185 191 GLU-RICH (ACIDIC).
CC SITE 1140 1144 GFFKR MOTIF.
CC DISULFID 72 83 BY SIMILARITY.
CC DISULFID 130 164 BY SIMILARITY.
CC DISULFID 698 754 BY SIMILARITY.
CC DISULFID 814 820 BY SIMILARITY.
CC DISULFID 884 898 BY SIMILARITY.
CC DISULFID 998 1023 BY SIMILARITY.
CC DISULFID 1031 1047 BY SIMILARITY.
CC CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 256 256 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 314 314 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 437 437 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 718 718 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 773 773 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 829 829 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 846 846 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 925 925 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 968 968 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1013 1013 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1086 1086 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1167 AA; 128983 MM; 88331C115DCCCFD CRC64;
Query Match 19.7%; Score 1157.5; DB 1; Length 1167;
Best Local Similarity 29.0%; Pred. No. 5.7e-69;
Matches 358; Conservative 211; Mismatches 459; Indels 205; Gaps 43;
QY 1 PNLDTENA--MTFQENARGFQSVVQLGSRVAVVVGAPQEIIVAAQNRGS-----LYQCDY 52
DB 20 FNMVDVWAVTALQAGAPAVLSSLLHLPDN-----NQTCLLVARRSNRNTALYRCAL 74
QY 53 STGSCPEIRLQVPEAVNMSGLSLAATT--SPQLLAC-GPTVHQTCSENYVKGICFL 109
DB 75 SI-SPDEIACQ-PVEHICWPGRGVGLTVGNHNGVLVCIQVQARKFRSLNSELTGACSL 132
QY 110 FGSNLRQOPQPEALRG-----C-----POE 131
DB 133 LTPNLDLQAAQYFSLGLEFLDPGAHVDSGDYCRSKGSGTGEEKSARRRRTVBEEDREED 192
QY 132 DSDIAFLIDGSGSIIPHRFBKMBPVSTVWEOQ--KSKTLFSLMOYSEEPRIHFTKEF 189
DB 193 GTEIAVLIDGSGSIIPHRFBKMBPVSTVWEOQ--KSKTLFSLMOYSEEPRIHFTKEF 252
QY 190 QNNPNRSLVRPIITOLLGRTHATGVRKVIKRELLMINTGARKNAFKLIVITGKFGDP 249
DB 253 RDINASLAKVQSIQVKEVKTASAMQVLDNIFIPSRGSRKKALKVWVLTDGDIQDP 312
QY 250 LGYEDVPEADBEGRVIRVIGVDAFRSEKSRQELNTIASPPRDHVFQVNFALKTIQ 309
DB 313 LNLITVINSPMQGVVRFAIGVDRFKNNITRELKLIASPKEMHTPKVINYALDGLL 372
QY 310 NQLEKIFAIEGTGTGSSSEHMSQGFSAITSNGP--LLSTVSGYDWAAGVFLY--TS 367
DB 373 SKLQQRIVHEMGT---VGDALQYQLAQTGFSAQILDKQVLLGTGVAFWNSGGLLYSTQ 429
QY 368 KEKSTFINMT-RVDS--DMNDAYLGVAAAILNRRVQSLVLAQPRYOHGLVAMERQVGM 425
DB 430 NRGCFNLQCTAKEDSRITVOYSYLGYSLAVLKAHGISYVAGAPRHURGAFTELKEDR- 488
QY 426 WESNA---NVKGTQICAFYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGQVQVCPLEP 482
DB 489 -EEDAFVRIEGQMGSYFGSVLCPVDIDMDGTTDFLLVAAPFYHIRGEGRVYVQVE 547
QY 483 GQARWQCDVAVYGEQGFQGRFGAALTVLGVNGDKLTDVAIGAP-----GEEDNRGA 536
DB 548 -QDASFSLSAHTLSGHFGTNSRFGFAMAAGVDINQDKFTDVAIGAPLEGFGAGDGASYS 606
QY 537 VYLFHGTSGSISPSHSGRIAGSKLSPLQYFGOSLSGQDLTMDGLVDLTWGAQGHVLL 596
DB 607 VYINGHSG-GLYDSFSQIIRASSVAGSLHFGMSVSGSLDFNGDLADIIVGSRDSAVV 665
QY 597 LRSQPLVAKVIMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKS-----TDRRLREG 653
DB 666 LRSRPVLDLTVSMFTF-----DALPMVPIGRM--DVNLCEVDSSVVASBFLGLEM 715
QY 654 QIQSVTVVDLALDSRPHRAVFNTRKSTRQTOVLGLTQC-----696
DB 716 FLNFTVDV-----TKORQLQEDSSGCGSLCKXNGSGSLCEHFWLI 760
QY 697 ETLKQLPNCIEDPVPVILNLFSLVGTPIPSAFGNLR-----PVLAEADAQLFTALP--P 750
DB 761 STEEL-----CEDCFPSNITIKVYE-----FOTSGGRDYENFTL--DHYKEPSAIFQLP 809
QY 751 FEKNGCNDNICODLSLITFSMSLCLVVGPPREFNVTTVVRNDEGDSYRTQVTFPPELD 810
DB 810 YEKQCKNVKFCIABIQLTIN--ISQOELVVGVTKEVTMMISLNSGDSYTMNMAINLPYN 868
QY 811 LGYRKVSTLQNRQSRMLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITPDV 870
DB 869 LQFKKI-----QKPVSPDVQCDKPKV---ASVLWACKIGHPII--KRSSVNVSVTWQL 918

QY 871 DSKASLGNKLLKANTVTSENNMPTNKTFQLELPVKYAVTMVTSHGVSSTKYLNFASE 930
DB 919 EESVFNPRADIVTITSNKSKEKSLAETR---SLOFHFHAFIADLSR--PSVMTWN--TSQ 971
QY 931 NTSRVNMQHQVSNLQGRSLPISLFLVPLVRLNQTVIMDRPQVTFSENLSST-----CHT 985
DB 972 SPDSHKFEFFNVHGEHLFGAVFQIQICVPIKLQDP-----QIVRVKNLTKTDHTECTQ 1025
QY 986 KERLPSSHDSPLAELRKAAPVNVCSIAVCQRIQCDIPFFGIQEEFNATLKGNSLDFDWIKTS 1045
DB 1026 SQSPAGSDSPVQVHKVSHVVCAL-----TSNKNVTVAEISVG 1065
QY 1046 HNHLIVSTA-----EILFNDVSFTLLPQGGAFVRSQETKVEFP-----EVNPLPLIV 1095
DB 1066 HTKQLLRDVSSELPIGLSIFSNKSLYELNAE-----NHRTKITVIFLKEETRSPLII 1119
QY 1096 GSSVGGILLIALITALYKLGFRFKROYKDWKSE 1128
DB 1120 GSSIGGLLVLLVIALIKCGFFKRYQQLNLZ 1152
RESULT 8
ITAE HUMAN
ID ITAE HUMAN STANDARD; PRT; 1179 AA.
AC P38570; Q9NZU9;
DT 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1
DE antigen) (CD103 antigen) (Integrin alpha-IEL).
GN ITGAE
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID:9606;
[1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.
RC TISSUE=Leukemia, and lymphocytes;
RX MEDLINE=94164962; PubMed=811947;
RA Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B.,
RA Parker C.M.;
RA "Molecular cloning of the human mucosal lymphocyte integrin alpha E
RA subunit. Unusual structure and restricted RNA distribution.";
RT J. Biol. Chem. 269:6016-6025(1994).
RN [2]
RP REVISIONS TO 88-114.
RA Parker C.M.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE OF 53-1179 FROM N.A.
RP TISSUE=Petal kidney;
RC MEDLINE=20138496; PubMed=10673275;
RA Touchman J.W., Anikster Y., Dietrich N.L., Maduro V.V., McDowell G.,
RA Stoteler-Suk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A.,
RA Green E.D.;
RA "The genomic region encompassing the nephropathic cystinosis gene
RT (CTNS): complete sequencing of a 200-kb segment and discovery of a
RT novel gene within the common cystinosis-causing deletion.";
RL Genome Res. 10:165-173(2000).
RN [4]
RN MUTAGENESIS OF ASP-109 AND PHE-316.
RP MEDLINE=20400502; PubMed=10837471;
RX Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y.,
RA Brenner M.B.;
RA "The role of alpha and beta chains in ligand recognition by beta 7
RT integrins.";
RL J. Biol. Chem. 275:25652-25664(2000).
CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
CC CELL MONOLAYERS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.

Db 757 SGSQCEBL-LIMPTEGELCEBDFCSNASVKVSYQL-QTPEGQTDHPQILDRYTEPFAI 814
Qy 747 ALPPEKXGNDNICODDLSTTFPMSLDCLVNGPREFNVTIVRNGEDSYETQTF 806
Db 815 FOLPYEKACKNKLFCVABQLA-TTVSQELVVGTLKELTILNINELTNGSDSYMTSWALN 873
Qy 807 FPLDLSYRKVSTLQNRQSRWRLACESASSTVSGALKSTSCSINHPIPPENSEVTFNI 866
Db 874 YERNLQ-----LKMQRKPPSPNIQCDPQV---ASVILNCRIGHPLV-LKRSSAHVS 923
Qy 867 TFDVDSKASLGNKLLKANVTSENN-----MPTNKTRFQ---LELPVKYAVYVWVTHSHV 919
Db 924 VMQLEENAFNPRTADITVTVNSNERRSLANETHLOFRHGFVAVLSKPSIMYNTQOGL 983
Qy 920 S--TKYLNFTASENTSRVMQHQYQVNSLQSRSLPISLVFLVPLVNLQTVINDRQPVTFSE 977
Db 984 SHHKEFLFVHGEN--LFGAEYQ-----LQICVPTKLGLOVAAVKKLTRQ 1028
Qy 978 NLSSTCHTKERLPHSDFLAELRKAPVNCISVNCQRIQCDIPFGIOBEFNATLKNLS 1037
Db 1029 ASTVCTWSQERACAYSS-VQHVEMHVSVCVIA-----SDKENVTVAAEIS 1073
Qy 1038 FDIWIKTSHNLLIVST-----ABILFNDVSFTLLPGQAGFVRSQVETKVEPPEVPNPL 1091
Db 1074 WD-----HSELLKDVTELOILGEISFNKSLYELGLAENH--RTKTVVFLKDEKYS 1125
Qy 1092 PLVNGSSVGGLLLLALITAAALKGLFPKQYKQMMSE 1128
Db 1126 PIIKSGVGLLVTLVILVLFKQGFPRKYQQLMLE 1162

RESULT 9
ITAL HUMAN
ID ITAL_HUMAN STANDARD; PRT; 1151 AA.
AC P56199;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a).
GN ITGA1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93155124; PubMed=8428973;
RA Briesewitz R., Epstein M.R., Marcantonio E.B.;
RT "Expression of native and truncated forms of the human integrin alpha
1 subunit.";
RL J. Biol. Chem. 268:2989-2996(1993).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-
E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD49a entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49a.htm".
DR PIR; A45226; A45226.
DR PDB; 1OC5; 17-MAY-00.
DR Genew; HGNC:6134; ITGA1.
DR MIM; 192968; -.
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; NAS.
DR GO; GO:0005518; F:collagen binding; TAS.
DR GO; GO:0007160; P:cell-matrix adhesion; NAS.
DR InterPro; IPR000413; Integrin_alpha.

InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3-
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00191; int_alpha; 5.
DR SMART; SM00327; VWFA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Repeat; Calcium; Magnesium; 3D-structure
PT DOMAIN; 1 1113 EXTRACELLULAR (POTENTIAL).
PT TRANSHEM 1114 1136 POTENTIAL.
PT DOMAIN 1137 1151 CYTOPLASMIC (POTENTIAL).
FT REPEAT 16 75 FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT DOMAIN 147 360 VWFA.
FT REPEAT 349 404 FG-GAP 3.
FT REPEAT 405 457 FG-GAP 4.
FT REPEAT 459 520 FG-GAP 5.
FT REPEAT 540 599 FG-GAP 6.
FT REPEAT 602 654 FG-GAP 7.
FT CA BIND 470 478 POTENTIAL.
FT CA BIND 552 560 POTENTIAL.
FT SITE 614 622 POTENTIAL.
FT SITE 1139 1142 GPEPR MOTIF.
FT DISULFID 54 64 BY SIMILARITY.
FT DISULFID 660 669 BY SIMILARITY.
FT DISULFID 675 728 BY SIMILARITY.
FT DISULFID 780 786 BY SIMILARITY.
FT DISULFID 850 858 BY SIMILARITY.
FT DISULFID 1002 1034 BY SIMILARITY.
FT DISULFID 1037 1044 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 855 855 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 880 880 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 887 887 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 946 946 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1151 AA; 127837 MW; 6B3F3C1AABF52808 CRC64;
Query Match 18.8%; Score 1102.5; DB 1; Length 1151;
Best Local Similarity 27.8%; Pred. No. 2.6e-65;
Matches 344; Conservative 211; Mismatches 487; Indels 195; Gaps 44;
Qy 1 FNLDENAMTFQENARG-FQSVVQL---QGSVVVVGAPQEIIVAANQRGSLYQCDYSTGS 56
Db 1 FNVVKNSMTFSGFVDMFGYTVQQYENESGKVLIGSLVVGQPKNRTGDVYKCPVGRGE 60
Qy 57 CEP-IRLOVPEA-----VNMSGLSLAATTSPPQLLAGPTVHQCSENYKGL 106
Db 61 SLPCVKLDLFWNTSIPNVTEVKENMTFGSL-VTPNNGGFLACGPLYAYRCGLHYTTGI 119
Qy 107 CFLFGSNLRQOPKQPFPEALRGCPQEDSDIAFLIDGSGSIIPHPF3RKMKEFVSTWELK- 165

Db 120 CSDVSPFTQVNSIAP--VQSCSTQ-LDIVVLDGNSIYFWD--SVTAFNLDLKRMXI 174
 Qy 166 -KSKTSLMOYSEFRIHFTFKFQNNPRSLVKPITQTLGR-THTATGVKRVIBELL 223
 Db 175 GPXQTVQVIGYGVNTHFENLKYSTEEVLAACKIVQGGQRTWALGTDTARKEAF 234
 Qy 224 NITGARNAPKILIVITDGEKFDPLGYEDVPEADREGVIRYVIGVDAFR-----SE 278
 Db 235 TEARGARGVKWVIVITDGEH-DNHELKKVQDCDENIQRFSAILOGSYNNGLSLSTE 293
 Qy 279 KSRQELNFIASPRDRHVQNNFEALKTIONQREKIPALEGTQTSSSSPEHEMSOEG 338
 Db 294 KFVEEIKSIASEPTKEKFFNVSDALVTIYKTLGERIFALEADQSAASFEXEMQTG 353
 Qy 339 FSAALTSGPLLSVGSYDAGGVFLYTSKE-----XSTF-INNRVDSMDNDAYLGAA 392
 Db 354 FSAHSDQWMLGAVGADWNGTVMQKASQIIPRNTTFNVSTKXNEPL-ASYLGVTV 412
 Qy 393 AILNRRVQSL-VIGAPYOHIGLVAMPRQNTGMSNANVKGQIQIGYFQASLCSVDVD 451
 Db 413 NSATASSGDVLYIAGQPRYNTGQVYIYRMDGNKILQTLGSGQIGSYFQSIILTTDID 472
 Qy 452 SNGSTDIVLIGAPHY-----YEQTR-GGQVSVCPPLPRQORARWQCDVAVLY 495
 Db 473 KDSMTDILLVGAPMNGTEKEEGQKVYVVALNQTREYQMLEPIKQCCSRQHNCTT 532
 Qy 496 GEQGPWG-RFGAALTVLGVNGDKLTDVAIGAPBEDNRGAIVLFGTSGSGISPHSQ 554
 Db 533 ENKEPOGARTGTAFAVKDLNDGENDIVIGAPLEDHGGAVTYHG-SGKTIRKEVAY 591
 Qy 555 RIASKLSPLRYQFQOSLGGDLTMDGLVDLTGVAQGHVLLRSQPLVRYKALWENFR 614
 Db 592 RPSGGDKTLKFFQGSIHGMDLMDGLDVTGIGLGGAAFWERDVAVVKTWNTFEN 651
 Qy 615 EVARNVFCNQVVKVKEAG--EVRVCLHVQ-KSTRDLRREGQIQSVVYTDLALDSGRPH 671
 Db 652 KVIQKQKCH--MEGKETVCINATCFEVLKSKEDTVEADLQ---YRVTLDLSLRQI 704
 Qy 672 SRAPVNET-----KXSTRQTOVLGLTQTCETLKLQPLNCIEPVPVILVPLNLSLWT 725
 Db 705 SRSPFGTQERKVORNIIVTRSEC-----TKHSFYMLDKHDFQDSVR---ITLDFNLTD 755
 Qy 756 PLSPFNGRLPVLAEADQRLFTALPFKNCGNDNICQDLSITTSFMSLCLVWGGPRE- 784
 Db 756 PENG-----PVLDDSLPNSVHEYPFADCGNKEKICISLSLHVATTEKDLIVRSQNDK 810
 Qy 785 FNVTVVRNDEGDSYRTQVTFPPLDLSYRKVSTLQNRORSRKLACESASSTEVSGAL 844
 Db 811 FNVSLTVKNTDSAYNTRTIHVHSPNLVFSGTEATQKD-----SCESN----- 853
 Qy 845 KSTSCSINHPIPPENSEVTENITDQVDSKASIGN-KLALKANVTSENNMPTNKTFOLE 903
 Db 854 HMITCKVGPFLRRGEMVTKLFPNTSYLNMENVTIVLSATSDSEEPPELSDNVNMS 913
 Qy 904 LPKVAVYVVTSHGVSTKYLNTASNTSRVMQHQYQVSN-----LGQRS-----L 950
 Db 914 IPVKYEVGLQFYS-SASEVHISIAANETVPEINSTEIDGNEINIFYLIRKSGFPMPDEL 972
 Qy 951 PLSLVP-----LVPRLNQTVIWRDPQVTFSENLSSCTHYE-----RLPS 991
 Db 973 KLSISFPNMTISNGYVLYPTGLSS-----SENANCRPHIFEDPFSINSKQMTT 1021
 Qy 992 HSDFLAELKAPWVCSTAVCQRIQDIPFFGIQE-----EFNATLK 1033
 Db 1022 STD---HLKRGTLQCNCKPATITCNLTSSDISQVNVSLILWKLTKFYSSNLNLTIR 1078
 Qy 1034 GNLSTFDWYIKTHNHLILVSTAEILFNDSVFTLLPQOGAFVRSQVETKVEFVENPLPL 1093
 Db 1079 GEL-----RSENASLVLSN-----QKRELAIQIKDGLPGRVPL 1114
 Qy 1094 --IVGSSVGLLILALITAAVLKGLFQKQYKDMWSE 1128

Db 1115 WVLISAPAGLLMLLILALWKIGFFKPLKKMKX 1151
 RESULT 10
 ID ITAH HUMAN STANDARD; PRT; 1189 AA.
 AC Q9UKK3; Q9UKQ1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-11 precursor.
 GN ITG11.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RS SEQUENCE FROM N.A.
 TT TISSUE=Petal heart, and Osteoblast;
 ED MEDLINE=99417678; PubMed=10486209;
 RA Lehnert K., Ni J., Leung E., Seijersen T., Gough S.M., Weaver A., Yao W.P., Liu D.,
 RA Wang S.-X., Morris C.M., Krissansen G.W.;
 RA "Cloning, sequence analysis, and chromosomal localization of the novel
 RT human integrin alpha11 subunit (ITG11).";
 RL J. Biol. Chem. 274:25735-25742(1999).
 RN [2]
 RS SEQUENCE FROM N.A.
 TT TISSUE=fetal muscle, and Uterus;
 ED MEDLINE=99395147; PubMed=10464311;
 RA Velling T., Kusche-Gullberg M., Seijersen T., Gullberg D.;
 RA "cDNA Cloning and Chromosomal Localization of Human alpha(11)
 RT integrin alpha-chain present in muscle tissues.";
 RL J. Biol. Chem. 274:25735-25742(1999).
 RN [3]
 RS SEQUENCE OF 954-1188 FROM N.A.
 TT TISSUE=Fibroblast;
 RA Andreu N., Estivill X., Baccarceller M., Sumoy L.;
 RA Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11
 CC -!- ASSOCIATES WITH BETA-1.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND
 CC HEART, INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO
 CC LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO
 CC FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS,
 CC SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN
 CC PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.
 CC -!- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING
 CC FETAL MUSCLE CELLS (IN VITRO).
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 1 VMFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AF109681; AAF01258.1; --
 CC EMBL; AF137378; AAD51919.2; --
 CC EMBL; AL359064; CAB94392.1; --
 CC HSSP; P17301; 1A0X
 CC Genew; HGNC:6136; ITG11.
 CC MIM; 604789; --
 CC GO; GO:0008305; C:integrin complex; TAS.
 CC GO; GO:0004895; P:cell adhesion receptor activity; TAS.


```

Db 1052 DLRRAPQLNHSNDVWSINCINRLVP-NQEIFHLLGNL---WLRSLKALKYKSMKIWN 1107
Qy 1053 STAETLFVDSVFTLLPGQGFVRSQTKVPPFVFN-----PLELVGSSVGGLLLLA 1106
Db 1108 AALGRQFH-SFP-----IFREDPSRQIVFISKQEDQVPIWIVGSLGGLLLA 1158
Qy 1107 LITLALYKLGFK--RQVDMMSGPGGAP 1136
Db 1159 LVLALWLGFGFRGARRR-----PGLDP 1183

```

RESULT 11

```

ITA2_BOVIN STANDARD; PRT; 1170 AA.
AC P53710;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GPIa)
GN ITGA2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94193647; PubMed=7511592;
RA Kamata T., Puzon W., Takada Y.;
RT "Identification of putative ligand binding sites within I domain of
RL integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).";
RJ J. Biol. Chem. 269:9659-9663(1994).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
CC COLLAGEN, INTEGRIN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-E-R IN
CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE
CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L25886; AAB59255.1; -
CC PIR; I45914; I45914.
CC HSSP; P17301; LAOX.
CC InterPro; IPR000413; Integrin_alpha.
CC InterPro; IPR002035; VWF_A.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; Integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC SMART; SM00191; Int_alpha; 5.
CC SMART; SM00327; VWA_1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS0234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.
FT NON_TER 1

```

```

FT SIGNAL <1 18
FT CHAIN 19 1170
FT DOMAIN 19 1121
FT TRANSMEM 1122 1143
FT DOMAIN 1144 1170
FT REPEAT 34 92
FT REPEAT 2 7
FT DOMAIN 177 367
FT REPEAT 2 7
FT REPEAT 423 475
FT REPEAT 477 538
FT REPEAT 540 599
FT REPEAT 604 656
FT CA_BIND 488 496
FT CA_BIND 552 560
FT CA_BIND 616 624
FT SITE 472 474
FT SITE 1146 1150
FT DISULFID 72 81
FT DISULFID 669 726
FT DISULFID 778 784
FT DISULFID 854 865
FT DISULFID 1008 1039
FT DISULFID 1044 1049
FT CARBOHYD 94 94
FT CARBOHYD 101 101
FT CARBOHYD 332 332
FT CARBOHYD 421 421
FT CARBOHYD 449 449
FT CARBOHYD 464 464
FT CARBOHYD 688 688
FT CARBOHYD 748 748
FT CARBOHYD 945 945
FT CARBOHYD 1063 1063
FT CARBOHYD 1070 1070
FT VARIANT 580 580
FT VARIANT 588 588
FT VARIANT 725 725
SQ SEQUENCE 1170 AA; 128929 MW; ECEFC15F2448FB1 CRC64;

Query March 18.5%; Score 1084; DB 1; Length 1170;
Best local Similarity 27.7%; Pred. No. 4.5e-64;
Matches 336; Conservative 216; Mismatches 495; Indels 168; Gaps 47;

Qy 1 FNLDTENATFQ-ENARGFGQSVQL---QGSRRVVVGAPOEIVAAORGSLYQC--DYST 54
Db 19 YNVGLPKATFSGPSSEQFGVAVQGFNPKGNWLLVGSFPGPKRNGDVYKCPVLDST 78
Qy 55 GCEPIRLQ-----VPVEAVMSLGLSLAATTSPPQLLACGPTVHTCSENVYKGLC 107
Db 79 TTCCKLNLQTSMSNVTEMKTNLSGLTLTRNVGTGGLTCGFLPKAQQCSQYTTGVC 138
Qy 108 FLFGSNLRQPKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFPVSTVMQLK-- 165
Db 139 SDVSPDP-QLRTSPAPAVQTCF-SFIDVVVVCDSNSIYPWD--AVKNFLKPVQGLDIG 194
Qy 166 KSKTLFSLMYSBFRHFTFKFQNNPNPRSLVKPTQLL---GRHTATGVKRVIRE 221
Db 195 PTKTQMGLIQVANNPRVFNLTNFKSKD---EMTKATSQTFQYGGDLNTFKAIQADRT 251
Qy 222 LNLITNGARKNAFKLIVITDGEKFGDPLGYEDVPEADREGVIRYIGV-----GDAFR 276
Db 252 AYSTAAGCPGATKVMVVVTVDGESH-DGSKLXAVIDQCNDKNILRFGLAVGLYLNALD 310
Qy 277 SEKGRQELNTIASKPPRDHVFQVNFNFEALKTIONLRKFIATGCTQTGSSSFHEHSQ 336
Db 311 TKNLIKETKAIASIPTRHFFNFVSDADLEKAGTIGEQIFSIETVQG-GDNFQMSQ 369
Qy 337 EGFSAAIT--SNGPLISTVGSVDWAGVFLYTSKEKSTFINMT--RVSDMN-DAYLGYA 391
Db 370 VGFSAEYSPQNNILMLGAVGAYDMSGTVQKTPHGLFLPSKQAEQIQLDRNHSYLGYS 429
Qy 392 AAIIILNRVQSLVLGAPRYQHIGLVAMPRQNTGMWESNANV-----KGTQIGVFGASL 445

```

```

Db 430 VASISTGNSVHFVACAPANTYQGVLYSVN-----ENGNTVTIQSQRGDQIGSYFSGVL 484
Qy 446 CSVDVDSGSDLVIGAPHYEQTR--GGQVSVCPLPRGORARQCDAVLYGEGQCPWG 503
Db 485 CAVDVNKDITDVLVAGAPMYNDLKKEBGRVYLFTIKG-ILNWH--QFLEGPNGLENA 541
Qy 504 REGAALTIVLGVNGDKLTDVAIGAPBEDNKGAVLFGHTSGSGISPSHSQRIAGS--KL 561
Db 542 RFGSAIALSDINMDGDFNDVIVGSPLENQNGAVYIYNGHEGM-IRLRYSKILGSDRAF 600
Qy 562 SPRLQVFGSISGGQDLTWGDLVDTVGAQGHVLLRSGQVFLRVKAIMFNPVREVARNVF 621
Db 601 SSHLYFSGSLGYDGLNGDSITDVSVCAGQVQVQLMSQSIADVSVDASFTPKXI--TLL 658
Qy 622 ECDQVVGKGEAGEVRVCLHVQKSTRDLRREGIQSVVYDIALD-----SORPHSAVFN 677
Db 659 KKNABSI-----KLKCF---SAKFRPTNQNNQVAIVYITIDQDFSRVSRIGLEK 707
Qy 678 EFGNSTRTQVLGLTQCE--TLKLQLPNCIEDPVSIVLRLEPFL--VGTPLSAFGL 733
Db 708 ENNERCLQKTHIVSAQRCSEVYIIHQPS-----DIISPLNLCMISLENPGT----- 756
Qy 734 RPLVAEDQRLTALFPPEKCGNDNICODLSITF-----SPMSLDCLVWGGPREFNTV 789
Db 757 NPALRAYSETKVFSGIPFKDGGDVCISDLVLNVQQLPATQQQPTIVSNQNKELTESV 816
Qy 790 FVRNDGDSYRQVTFEFLDLSYRKVSTLQNRQSRWRLACESAST-EVSGALKSTS 848
Db 817 QLNKKEGAYNTEIVVDSENLFF-----ASKMPVDGTEVTCQIASQKSVT 864
Qy 849 CSINHPITPENSEVFTNITPDVDSKASLGNKLLKANVTSNNMPTNKTFFQLELPVKY 908
Db 865 CNVGYPALKSQKVTFITNFDPLNQ-NLQNASISFRALSQSEENMADNSVNLKLSLLY 923
Qy 909 AYVMVTVSGVSTKYLNTASNTSRVWQHQQVQVSNLQOR-----SLPISLFLV 958
Db 924 DAEIHIT-RSTNINPYEVLSDGNVSSV-HSFE--DIGKPFIFSIVTGVSPVMSA--- 976
Qy 959 PVRLNQTVWRPQVTFSEN---LSSTCHTKE-----RLPSHSDFLAE- 998
Db 977 -----SVLIHPQVTKDKNPLWLTGVHTDQAGDISCEABINPLKIGQTSVSSVFSKEN 1030
Qy 999 LRKAPVNCISAVCQRIQCDIPFGIQEFPNATLKNLSFDWYIKTSNHLIIVSTARI- 1057
Db 1031 FFIKELNCRATSCSNIMCWLEDLQVGEYFLNVSTRIWNGTFAASTFQTVQLTAAAEID 1090
Qy 1058 LFNDVSFTL-----LPGQAFVRQVETKVEPFE-VENPLPLVSGSVGGLLIALITA 1110
Db 1091 TYNPQIYVIENTVIP-----LTINKPHEKVEVPTGVIVGVIAGILLALLVA 1140
Qy 1111 ALYKLGFFKRYQKDM 1125
Db 1141 ILWKLGFRRKRYERM 1155

RESULT 12
ITR2_MOUSE
ID ITR2_MOUSE STANDARD; PRT: 1178 AA.
AC Q62469; Q62163;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 x C5A; TISSUE=Lung;

```

```

RX MEDLINE=94363406; PubMed=8081889;
RA Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,
RA Danjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.;
RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but
RT not virus binding.";
RL Cell Adhes. Commun. 2:131-143(1994).
RN [2]
RN SEQUENCE OF 450-1178 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94355691; PubMed=7521231;
RA Wu J.B., Santoro S.A.;
RT "Complex patterns of expression suggest extensive roles for the alpha
RT 2 beta 1 integrin in murine development.";
RL Dev. Dyn. 199:292-314(1994).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING
CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
CC COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN
CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION
CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL; Z29987; CAAS2877.1; -.
CC EMBL; X75427; CAAS3178.1; -.
CC PIR; S44142; S44142.
CC HSP; P17301; IAOX.
CC MGD; MGI:96600; Itga2.
CC InterPro; IPR000413; Integrin_alpha.
CC InterPro; IPR002035; VWFA.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; Integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC SMART; SM00191; Int_alpha; 4.
CC SMART; SM00327; VWFA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS00234; VWFA; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Platelet; Signal; Repeat; Calcium; Magnesium.
CC SIGNAL 1 26
CC CHAIN 27 1178
CC DOMAIN 27 1129
CC TRANSMEM 1130 1151
CC DOMAIN 1152 1178
CC REPEAT 42 100
CC REPEAT ? ?
CC DOMAIN 185 375
CC REPEAT ? ?
CC REPEAT 431 483
CC REPEAT 485 546
CC REPEAT 548 607
CC REPEAT 612 664
CC REPEAT 496 504
CC CA_BIND 560 568
CC CA_BIND 624 632
CC SITE 480 482
CC SITE 1154 1158
CC DISULFID 80 89
CC BY SIMILARITY.
CC CELL ATTACHMENT SITE (POTENTIAL).
CC GFGR MOTIF.
CC BY SIMILARITY.

```

```

FT DISULFID 677 734 BY SIMILARITY.
FT DISULFID 786 792 BY SIMILARITY.
FT DISULFID 862 873 BY SIMILARITY.
FT DISULFID 1016 1047 BY SIMILARITY.
FT DISULFID 1052 1057 BY SIMILARITY.
FT CARBOHYD 102 102 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 109 109 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 429 429 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 472 472 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 696 696 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1054 1054 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1178 AA; 128926 MW; 1F194B9C0240F465 CRC64;

Query Match
Best Local Similarity 28.1%; Score 1072; DB 1; Length 1178;
Matches 344; Conservative 206; Mismatches 487; Indels 188; Gaps 44;

QY 1 FNLDTENMTFO--ENARGGQSVQI---CGSRVVVGAFOEIVANORGLYQC--DYST 54
Dy 27 YNVLFGAKIFGSPSEQGYVQVLTNPQGNWLLVSPWSPFENRMGDVYKCPVDLPT 86
QY 55 GSCERLQ-----VPVEAVMSLGLSLAATSPQLLAGPTVHQCSENIVKGLC 107
Dy 87 ATCEKLANQNSASISNTEIKTMSLGLTLTNPTGGTGLTQPLWHQCGNQYATGIC 146
QY 108 FLFGSLRQOPQ---KFPALRGCPQEDSDIAFLIDGSGIIPHDPFRMKFVSTVMEOL 164
Dy 147 ---SDVSPDFQPLTSFSAVQACPSL--VDVVVWCDESNIYP--WEAVKFLVKFVTGL 199
QY 165 K--KSTLSLMOYSEPIHFTFKFQNNPRSLVKPITQLLG--RTHATGVRKVIRE 221
Dy 200 DTGPKKTQVALIQYANEPRIIFNLNDFETKEDMVOATSETROGGDLTNTFRAIFARDY 259
QY 222 LNNITNGAKNAFKILIVITDGEKFGDPLGYEDVPEADREGVIRVIGV-----GDAFR 276
Dy 260 AYSQTSGGPRGATKVMVVYTGESH--DGSKLTVIQCCNDDEILRGIAVLGLVNRALD 318
QY 277 SEKSRELNTIASKPRDRHVQNNFEALKTQNLQREKIPAEIHGTQCGSSSFHEMSQ 336
Dy 319 TKNLKEIKALASTPTERYFFNVADEALALKAGTLGQIFSGIEGTVOG--GDNFQKEXAQ 377
QY 337 EGPSA--AITSNGPLASTVGSYDWAGSVFLYTSKEXSTFINMT--RVDSDMN--DAYLGYA 391
Dy 378 VGFSAIYAFQNDILMGAAGAPWMSGTIVQFETSHKVIPEPKQAFQVLDNRHESFLYS 437
QY 392 AAILIRNVQSLVLAGPYQHIGLVAMFRQNTGWSNANV-----KGTQIGAYFGASLCS 447
Dy 438 VAAISTEDGVHPVAGAPRANYTCQIVLYSVNK---QGNVTVIQSHRGDQIGSYFGSVLCS 494
QY 448 VDVSNGSTDLVLIGAPHYEQTR--GGQVSVCLPRGQARQCCDAVLVYGGQGPWGRF 505
Dy 495 VDVKDITITDVLVIGAPTWNLDKKEGKVLFTITKGLINHQH----FLEGPEGTGNARF 551
QY 506 GAALTVLGVDVNGDKLTDVAIGPEEDNRGAVLPHGTSGSITSPSHSQRIAGSKLSPR- 564
Dy 552 GSAIAALSDINMDGENDVIVGSPVENENSGAVIYNGHQT--IRTKYSQKILGNSGAFR 610
QY 565 -LQYFGQSLSGGDLTDGLVDLTGCAQGVLLRSOPVLRYKAINFENPREVARNVFEC 623
Dy 611 HLQFFGRSLGVDLNGDSITDYSIGALGVQIQLWSQSADVAIEALFTF----- 660
QY 624 NDQWVGKEAGEVRVCLHVQKSRDLRREGQIQSVVTVYDALD-----SCRPHSRVFNFT 679
Dy 661 -DKITILNKDAKITLKCFAEFAPAGQNNQV--AILFNMTLDAGHSSSRVTSRGVREN 717
QY 680 KNSTRQGVQLGTQTCET--LKLQLPNCIEDVPSIVILRNPLSVGTPLSFGNLRPVL 737
Dy 718 SERFLOKVMVNEVQKSEHHISIQKPS---DVNPLDLRVDISLENPGTS-----PAL 768
QY 738 ABDAQRLFTALFPFECNCGNDNICQDDLSI-----TFSPMSLDCVLVGVGPREFNT 788

```

RESULT 13

```

IT9A2 HUMAN
ID IT9A2_HUMAN STANDARD; PRT; 1181 AA.
AC P17301;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITGA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.
RC TISSUE=Endothelial cells;
RX MEDLINE=89308879; PubMed=2545729;
RA Takada Y., Hemler M.E.;
RA "The primary structure of the VLA-2/collagen receptor alpha 2 subunit
RT (platelet GP1a): homology to other integrins and the presence of a
RT possible collagen-binding domain.";
RL J. Cell Biol. 109:397-407 (1989).
RN [2]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,
RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
RL Submitted (May 2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.
RX MEDLINE=98019223; PubMed=9353312;
RA Emsley J., King S.L., Bergelson J.M., Liddington R.C.;
RT "Crystal structure of the I domain from integrin alpha2beta1.";
RL J. Biol. Chem. 272:28512-28517 (1997).
RN [4]
RP VARIANT HPA-5 (BR).
RX MEDLINE=94043762; PubMed=7901236;
RA Santoso S., Kaib R., Walka M., Kiefel V., Mueller-Eckhardt C.,
RA Newman P.J.;
RT "The human platelet alloantigen Br(a) and Brb are associated with a
RT single amino acid polymorphism on glycoprotein Ia (integrin subunit

```


Db 90 ATCEKLNQSTSPNVTETKMSLGLILTRNMGTCGLTCPLMAQCCGQNYTTTVC 149
Qy 108 FLFGNLRQOPKPEALRGCPQSDSDIAFLIDGSGIIPHDPRMKEPFTVWQK-- 165
Db 150 SDISPDF-QLSASPSAPQPCPSL-IDVVVCDSDSNIYWD--AVKNEFKFVQGLDIG 205
Qy 166 KSKTLFSLMOYSBEPRIHFTFKBNPNPRLSKVPKIPOLLG-RTHATGTVKRVIRELLN 224
Db 206 PTKQVGLIQIANPRVFNLTNTYKTEEMIVATSCSYQSGDLTNFTGAIQYARKAYS 265
Qy 225 ITGARKNAFKILIVITDGEKFGDPLGYEDVIPRADREGVIRYVIGV-----GDAFRSEK 279
Db 266 AASGRRRATQVMVVTVDGSH-DGSMKKAVIDOCNHDNIRLFGIYVLYNRLNALTQY 324
Qy 280 SQELNTIASPPRDHVFQVNFALKTIQOLREKIPAIESTGTGSSSSPEHMSQEGF 339
Db 325 LIKEIKATASIPTEYFPNVNDEALEKAGTLGEGHIFSIETVQG-CDNFQEMWSQVGP 383
Qy 340 SAAITSNGP--LLSTVGYSDMAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYIGYAAAI 394
Db 384 SADYSSQNDILMLGAVGAGSGTIVOKTSHGLIFPKQAFQDILQDRNHSYLGYSVAA 443
Qy 395 ILNRNVQSLVIGAPRYQHILGVAMFRONTGWSNANV-----KGTQIGAYFGASLCSV 448
Db 444 ISTGESTHFAVAGAPRANTYQGVLYSVN-----ENGNITVIOAHRGQDQIGSYFGSLCSV 498
Qy 449 DVDNSGSDLVILGAPHYETQTR--GGQVSVCPLEPRGQARWQCDVAVLYGEGQGFWRGFG 506
Db 499 DVDKDTITDVLVAGPMYSDLKKEEGRYLFTIKKGILGQH-----FLEGPEGIENTRFG 555
Qy 507 AALTVLGVNGDKLTDVAGPESDNRGANVLFHGTSGSGISPSHSORIASG--KLSPR 564
Db 556 SAIAALSINMDGDFNDVIGSPLENQSGAVYIYNGHQT-IRTKYSQKILGSDGAPRSH 614
Qy 565 LOYFGQSLGGODLMDGLVDLTVCAQGHVLLLRQPVLRVKAIMEFPREVARNVFECN 624
Db 615 LQYFGRSLDGYDLGDSITDVSIGAFQGVQVQNSQADVAIBASFTPEKI--TLVNKN 672
Qy 625 DQVWKGKAGEVRVCLHVQKSTRDLREQIQSVVTVYDLAD-----SGRPSRAVFNETK 680
Db 673 AQII-----LKLCF-----SAKFRPTKQNNQVAIVAITLDDADGSSRVTSGFLKKN 721
Qy 681 NSTRAQTVLGLTQTC--ATLQKLPNCIEDPVSPIVLRNLNSLAVGTPLSAPGNLRPVL 738
Db 722 ERLQKQVMVNAQSCPEHIIYQEPS-----DVNSLDLRAVDLSLENPGTS-----PALE 772
Qy 739 EDAQRLFTALFFPEKNCNDNICODLSITP-----SPMSLCLVGVGPRFNVTVVRND 794
Db 773 AYSETAKVSIIPFKDCGDLGCLISDLVDVQIIPAAQOPRIVSNQNKRLTFSVTLNKK 832
Qy 795 GEDSVRTQVTFPPLDLSTKRVSTLQNSQSWRLACESAST-EVSGALKSTSCSINH 853
Db 833 RESAYNTGIVDFSENLPF-----ASFSLPVDGTEVTCQVAASQKSVACDVGT 880
Qy 854 PIPPENSEVTNITFDVDSKASIGNKLLKANVTSENMPRTNKTFFQLELPVKYAVVMV 913
Db 881 PALKREQVTFINFDNLQ-NLQNASLSFQALSSEQENKADNLVNLKILLYDAEI- 938
Qy 914 VTSHGVTKYLNAPTASENTSRVMOHYQVSNLQOR-----SLPSLVELV----- 958
Db 939 ---HLTRSTNINPFISSDGNVPSIVHSFEDVGPKFIFSLKVTGSGVPVSMATVIHIPQ 995
Qy 959 -----PVLNQTVIWRDPOVTP-SENLSSTCHTKER 988
Db 996 YTKERPLMYLTVQVTDKAGDISCNADINPLKIGQT-----SSSVSPKSENER---HTXE- 1047
Qy 989 LPSHSDPLAELRKAPVNVNCSIAVCORIQCDIPFGIOEFNATLKNLSDFWYIKTSHNH 1048
Db 1048 -----LNCRTASCSNVTCWLDKDHMKGEYFVNVTVTRINWNGTFASSTFQT 1091
Qy 1049 LLIVSTAEI-LFNDVSFTLLPGQAFVRSQTEKTVBEFVENP-----LP--LIVGSSVG 1100
Db 1092 VQLTAAAEINTYNPFIYVI-----EDNTVTIPLIMKDEKAEVPTGVIGSIITA 1141

Qy 1101 GILLALITAAALYKLGFFKROVKDM 1125
Db 1142 GILLALVALVALLKLGFFKRYEKY 1166
RESULT 14
ITAG HUMAN
ID ITAG HUMAN STANDARD; PRT; 1167 AA.
AC 075578; Q9UH28;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-10 precursor.
GN ITGA10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Articular chondrocytes;
RX MEDLINE=98352078; PubMed=9685391;
RA Camper L., Hellman U., Lundgren-Akerlund E.;
RT "Isolation, cloning, and sequence analysis of the integrin subunit
RT alpha10, a beta1-associated collagen binding integrin expressed on
RT chondrocytes.";
RL J. Biol. Chem. 273:20383-20389(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Endothelial cells, and Heart;
RX MEDLINE=20169197; PubMed=10702680;
RA Lehnert K., Ni J., Leung E., Gough S.M., Morris C.M., Liu D.,
RA Wang S.-X., Langley R., Krissansen G.W.;
RT "The integrin alpha10 subunit: expression pattern, partial gene
RT structure, and chromosomal localization.";
RL Cytogenet. Cell Genet. 87:238-244(1999).
CC 1- FUNCTION: INTEGRIN ALPHA-10/BETA-1 IS A RECEPTOR FOR COLLAGEN.
CC 1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-10
CC 1- ASSOCIATES WITH BETA-1.
CC 1- SUBCELLULAR LOCATION: Type I membrane protein.
CC 1- TISSUE SPECIFICITY: Widely expressed with highest expression in
CC muscle and heart. Found in articular cartilage.
CC 1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC 1- SIMILARITY: Belongs to the integrin alpha chain family.
CC 1- SIMILARITY: Contains 1 VWFA domain.
CC 1- SIMILARITY: Contains 7 FG-GAP repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; AF074015; AAC31952.1; -;
DR EMBL; AF112345; AAF21944.1; -;
DR EMBL; AF172723; AAF61638.1; -;
DR HSSP; F17301; IAOX.
DR Genew; HGNC:6135; ITGA10.
DR MIM; 604042; -;
DR GO; GO:0008305; C: integrin complex; TAS.
DR GO; GO:0004895; F: cell adhesion receptor activity; TAS.
DR GO; GO:0005518; F: collagen binding; TAS.
DR GO; GO:0007160; P: cell-matrix adhesion; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRQ1185; INTEGRINA.
DR PRINTS; PRQ0453; VWFADOMAIN.

DR SMART; SM00191; Int_alpha; 4.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.
 DR PROSITE; PS00344; VWA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Calcium; Magnesium.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 1167 INTEGRIN ALPHA-10.
 FT DOMAIN 23 1122 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1123 1145 POTENTIAL.
 FT DOMAIN 1146 1167 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 38 97 FG-GAP 1.
 FT REPEAT ? 9 FG-GAP 2.
 FT DOMAIN 167 350 VWA.
 FT REPEAT 365 427 FG-GAP 3.
 FT REPEAT 428 482 FG-GAP 4.
 FT REPEAT 483 545 FG-GAP 5.
 FT REPEAT 546 605 FG-GAP 6.
 FT REPEAT 608 660 FG-GAP 7.
 FT DOMAIN 1134 1140 POLY-LEU.
 FT CA_BIND 494 502 POTENTIAL.
 FT CA_BIND 558 566 POTENTIAL.
 FT DISULFID 76 86 BY SIMILARITY.
 FT DISULFID 666 675 BY SIMILARITY.
 FT DISULFID 691 736 BY SIMILARITY.
 FT CARBOHYD 799 795 BY SIMILARITY.
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 733 733 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 763 763 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 921 921 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1011 1011 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 844 844 I -> L (IN REF. 2).
 FT CONFLICT 909 909 G -> V (IN REF. 2).
 FT CONFLICT 926 926 E -> D (IN REF. 2).
 SQ SEQUENCE 1167 AA; 127573 MW; A67D3A1C25C1ABAO CRC64;

Query Match
 Best Local Similarity 18.1%; Score 1065.5; DB 1; Length 1167;
 Matches 358; Conservative 196; Mismatches 488; Indels 181; Gaps 46;
 1 FNLDENAMTFQENARG-FQGVVQLQSGR---VVVGAPQEIIVAAQNGSLYQC----- 50
 23 FNLDHHPRLFPQPPAEFGYSVLQHVGGQRMVLVGA PWGPGSDRRGDVYRCPVGGAH 82
 51 -----DYSTG-SCEPIRLQVPVAVNMSLGLSLAATSPPLLACGTVHTQCS 99
 83 NAPCAKGLLDYQLGNSSHP-----AVNMHLGMSLLETGDDGFMACAPLWSRAGS 134
 100 NTYVKGCLFLPGSNLRQQPKPEALRGCPQEDSDIAPLDGSGSIIPHD-----FRRM- 153
 135 SVFSSGICARVDASFPQGS LAPTAQR-CETY-MDVVIVLDGNSIYPMSEVQTFLRLV 192
 154 -KEFVSTVMEQLKKSKTLFSLMYSREPIHFTTFKEFQNNPNRSLVKDITQLGR-THT 211
 193 GKLFIDP--EQIQ-----VGLVQYGESPVHVESLGDFTKKEVYRAAKNLSRREGRTKT 245
 212 ATGVKRVIRELLNITGARKNAFKILVITDGKFE-GDPLGYEDVIPEDRBCGVIRYVIG 270
 246 AQAIMVACTEGFSQSGHGRPEARLLVVDVGDHSGHEELPAALKAACEAGR--VTRYGIA 303
 271 V-GDAFSEKS-----RQELNTIASKPPDRHVFVQNNFEALKTIQNLQREKIFAIBQTQ 325
 304 VLGHYLRQRDPSPFLREIRTIASDPDRFPFNVDTAAALTDIVDALGDRIFGLESHAB 363
 326 SSSSFHEMSQEGFSAITSNGELLSTVGSYDHAGGVFLYTSKEKSTFFINWTVDS---- 381

364 NESSPGLMSQIGFSTHRLKDGILFGVGVAYDWGGSVLMLEGGHRLFPFPRMALEDFPPA 423
 382 -DNNDAYLGYA-AAIILNENRQSLVLGAPRYOHIGLVAMFR-ONTGWNESNANVKGQIG 438
 424 LQNHAYLYGYSVSSMLLRGGRRLFLSGAPFRERGVIAFQPKDKGAVRVAQSLQGHQIG 483
 439 AYFCASLCSVDVDSNGSTDLVLGAPHY--EOTRGQVSVCPLEPRGQORAWQCDVLYG 496
 484 SYFGSELPLDTRDGTDTLLVAAPMFLGPONKTRGVYVILV--GQOELLTLQGTLP 541
 497 EQGQWGRFGAALTVLGDVNGDKLTOVAIGAPGEENRGAVLFGHTSGSGISPSHSORI 556
 542 EPPQD-ARFGFAMGALPDLNQGFADVAVGAPLEDGEGQALYLYHGTO-SGVAPHAPRI 599
 557 AGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGEVLLRSQPLRVKAIMFNPREV 616
 500 AASMPHALSYFGRSVDRLDGLDDLDVAVGAQGAAILLSRSPVHLTSPLEVTPOAI 659
 517 ARNVFECNDQVVKRAG--EVRVCLHVQKSTRDLREGQIOSWVYDIALDSCGRHSRA 674
 660 SVVQRDCRR--RQGEAVCLTAALCFQVTSRTGWRDH---QFYERFTASDEWTAGARA 713
 675 VFNET--KNSTRQTVLGLTQTCETLKLQLPNCIEDPSPVIVLRNLFSLVGTPLSPAGN 732
 714 AFDGSGQRLSPRLRLSVG-NVICEQLHFHVLDTSDYLRPVALTVTFALDNTTKPG-- 768
 733 LRPVLAEDAQRLFTALPPPKNGCNDNICODDLISITFSMSLDC-----LVVGGPRE 784
 769 --PVLNBSPTSIOKLVPFSKCGPDNECVTLVLQ---VNMDIRGSRKAPFVVRGGRKK 823
 785 FNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWFLACESASSTEVSAL 844
 824 VLVSTTIENKENAYNTLSIISRNL---ELASLTQPR-ESPIKVECAAPSA----- 872
 845 KSTSCSINHPIPPENSEVTNITPDVDSKASLG---NKLL-----LKANVTSENMPRTN 896
 873 HARLCSVGHVPVQTAKVTLLEPEFSCSLLSQVPGKLTASSDSLRNGTLQENTAQT- 931
 897 KTEQLZLPVYAVYVMTSHGVSTKYNLFTASNTSRVMQHOY-----Q 941
 932 -----SAYIQVEPH-----LFSSESTLHYEVHPVCTLPVGPPEKTLR 973
 942 VSNLG---QRSPLISLVFLVP-----VRLNQTVMWRPQVTFSENLSSTCHTKERL 989
 974 VQNLGCVTVSGLIISA--LLPAVAHGNYFLLSQVI-----TNNASCIQNLTEP 1022
 990 PSHGDFLAELRKAPVNVNCIAVCQICDIPFGIQEFNATLKNLSFDWVYKTSNHL 1049
 1023 PGPPVHPEELQHTNRLNGSNTQCVVRCHLQQLAKGTEVSGLLRLVINEFFRAKPKSL 1082
 1050 LIVSTAEILFNDSVFTLLPGQAFVRSQETKTVKFFEPVNPPLIVGVSSVGLLILALIT 1109
 1093 TVSTFELGTEGSGVQLQTEASRWSLLEV-VQTRPILISLWILIGSLVGLGLLALLV 1141
 1110 AALYKLGFF-----KRYK 1123
 1142 FCLWKLGFPAHKKIPEEKREEK 1164
 RESULT 15
 ID ITAL RAT STANDARD; PRT; 1180 AA.
 AC P18614;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Integrin alpha-1 precursor (Laminin and collagen receptor) (VLA-1)
 DE (CD49a).
 GN ITGAI.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;

[1]
SEQUENCE FROM N.A.
RA MEDLINE=90338125; PubMed=2380249;
RA Ignatius M.J., Large T.H., Houde M., Tawil J.W., Barton A.,
RA Esch F., Carbonetto S., Reichardt L.F.,
RT "Molecular cloning of the rat integrin alpha 1-subunit: a receptor
for laminin and collagen.";
RL J. Cell Biol. 111:709-720(1990).
[2]
X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 151-364.
RX MEDLINE=99313197; PubMed=10386626;
RA Nolte M., Pepinsky R.B., Vanyaminov S.Y., Koteliarsky V.,
RA Gotwals P.J., Karpus M.,
RT "Crystal structure of the alpha1beta1 integrin I-domain: insights into
integrin I-domain function.";
RL FEBS Lett. 452:379-385(1999).
CC !- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-
E-R IN COLLAGEN.
CC !- SUBUNIT: HETERO-DIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
ASSOCIATES WITH BETA-1.
CC !- SUBCELLULAR LOCATION: Type I membrane protein.
CC !- DOMAIN: THE INTEGRIN I-DOMAIN (INSET) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC !- SIMILARITY: Belongs to the integrin alpha chain family.
CC !- SIMILARITY: Contains 1 VWFA domain.
CC !- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
DR EMBL; X52140; CAA36384.1; -
DR PIR; A35854; A35854.
DR PDB; 1CK4; 03-MAY-00.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWFA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Calcium; Magnesium; 3D-structure.
FT SIGNAL 1 28 INTEGRIN ALPHA-1.
FT CHAIN 29 1180 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 29 1142 POTENTIAL.
FT TRANSKEM 1143 1165 POTENTIAL.
FT DOMAIN 1166 1180 CYTOPLASMIC (POTENTIAL).
FT REPEAT 44 103 FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT DOMAIN 175 388 VWFA.
FT REPEAT 377 432 FG-GAP 3.
FT REPEAT 433 484 FG-GAP 4.
FT REPEAT 485 565 FG-GAP 5.
FT REPEAT 567 626 FG-GAP 6.
FT REPEAT 629 681 FG-GAP 7.
FT CA BIND 497 505 POTENTIAL.
FT CA BIND 579 587 POTENTIAL.
FT CA BIND 641 649 POTENTIAL.
FT SITE 1168 1172 GFFKR MOTIF.
FT DISULFID 82 92 BY SIMILARITY.
FT DISULFID 687 696 BY SIMILARITY.
FT DISULFID 702 755 BY SIMILARITY.
FT DISULFID 807 813 BY SIMILARITY.
FT DISULFID 877 885 BY SIMILARITY.
FT DISULFID 1029 1062 BY SIMILARITY.

FT DISULFID 1066 1073 BY SIMILARITY.
FT CARBOHYD 100 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 105 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 112 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 217 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 317 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 341 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 418 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 531 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 698 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 747 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 779 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 820 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 839 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 882 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 907 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 938 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 965 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 973 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1007 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1084 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1103 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1114 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1180 AA; 130808 MW; 8E5DA2BE02362EE4 CRC64;
Query Match 18.0%; Score 1060; DB 1; Length 1180;
Best Local Similarity 27.4%; Pred. No. 1.8e-62;
Matches 344; Conservative 201; Mismatches 478; Indels 234; Gaps 47;
QY 1 FNLDTENAMTFOENARG-FGQSVQVL---QSRVVGAPQBIIVANQSGISYQCDYSTGS 56
DB 29 FNVVDKNSMSPGVEDNFGYTVQOYENEGKWLIGSLVPGQPKARTGDVVKCFVGR 88
QY 57 CEP-IRLOPVVEA-----VMSLGLSLAATSPPOLLAGCPVTGHTCSEYTVKGL 106
DB 89 AMPCVKLDLPVNTSIPNTEIKENMTFGSTL-VTNPGFGFLACGLYAYRCGLHYTTGI 147
QY 107 CFLFGSNLRQOPKPEALRCPQEDSDIAFLIDGSGIIIPDFRMEKFVSTVMEQLK- 165
DB 148 CSDVSPTEQVWNSPAP--VQECSTQ-LDIVIVLDGNSIYP--MESVIAFLNLLKRM 202
QY 166 -KSKTLPSLMQYSEPRTHPTFKPEQNNPRLSVKPTQLLG-RTHATATGVKVI 223
DB 203 GPKQTQGVQVGENVTHFNLNKYSTSTEEVVAANKIGRQGLGLTMTALGIDTARKEAF 262
QY 224 NITNGARKNAFKILIVITDGBKFGDPLGVEDVPEADREGVIRYVIGVDAFR-----SE 278
DB 263 TEARGARGVKVMVIVTDGESH-DNYRLKQVIOQCEDENIQRFSAIILGHYNRGNLS 321
QY 279 KSRQELINTASKPPRDHVQVNNFALKTINQLREKIPATBGTCTGSSSSEHEMSORG 338
DB 322 KFEVEBIKSIASEPTEKHPFNVSDELALVTIVKALGERIFALEATADQSAASEMMSQTG 381
QY 339 PSAALTSNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSMDND---AYLG 393
DB 382 FSAHYSQDWMLGAVGADWNGTVVMQKANKVIPHNTTFQTEPAKMEPLASLYGT 441
QY 394 IILNRVQSLVGLGAPRYOHIGLVAMFRONTGMWESNANVKGTOICAYFGASLCSDVD 453
DB 442 SATIPGDVLYIAGPRYNHTGQVWVYQWEDGNINILQTLGGEGQISYFGSVLT 501
QY 454 GSTDLVLGAPHY-----YEQTR-GQVSVCPPLRQQRARWCDVLYGE 497
DB 502 SYTDLLVGVPMYMGTEKEEQQKVYVAVNQTRFQYQNSLEPIRQTCSSSLKDN 561
QY 498 QGQPMG-RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFGTSGSGISPSHQ 556
DB 562 KNEPCGARFGSTAIAAVKDLNVDFDNDVVICAPLEDHAGAVYIFG-SGKTREAYAO 620
QY 557 AGSKLSPLOYGQSLSGQDLTMDGLVLTVCAGHVLVLLSQPVLRYKAIEMFPREV 616

Db 621 PSGBGKTLKFFGQSIHGEYDNGDGLTDTVTIGGLGGAALFWARDVAVVVKVTMTNFPNKV 680

QY 617 ARNVFECNDQVVKGEAG--EVRVCLHQV-KSTRDLREGQIQSVVTVYDLALDSGRPHSR 673

Db 681 NIQKNCR---VEGKETVCINATMCFHVVKLSKEDSIVEADLQ-----YRVTLDSLRQISR 733

QY 674 AVFNET-----KSTRRTQVGLTQTCETLKLQLPNCI-----EDPVSPIVLR 718

Db 734 SPFSGTQEKIORNITVRESE-----CIRHSFYWLKDKHDFQDSVRVTL 776

QY 719 NFSLVGTPLSAFGNLKAPVLAEDAQRFLTALPFPEKNCNDNICQDDLSITFSFNSLDCIV 778

Db 777 DENLT-DPENG-----PVLDDALPSVHSHIFPAKDCGNKERCISDLTLNVSTTEKSILLI 830

QY 779 VGGPRE-FNVTVVRNDGSDSYRTQVTFPFLDLSYRKVSTLQONORSQBSWGLACBSASS 837

Db 831 VKSCHDKFNVSUVRKNGDSAYNTRTVQVHSPNLIFSGIEEIQD-----SCSN-- 880

QY 838 TEVSGALKSTCSINHPIFPENSEVTFNITFDVDSKASLGNKLL-LKANVTSENMPRTN 896

Db 881 -----QNTCRVGYFFLRAGETVTKIIFQFNTSHLSENALIHLSATSDSEEPLESIN 933

QY 897 KTEFQLELPPVKYAV-----YXVVTSHGVST-----KYLNFTASNTSRVMOHQYQVSNL 945

Db 934 DNEVNISSIPVKEVGLQFYSSASEHHISVAANETIPEFINST--EDIGNEINNVFTIRKR 991

QY 946 QOESLP---ISLVP-----LVPVRLNCTVWD-----RP-----Q 972

Db 992 GHFPMPELQISIFPNLTADGYPVLPIG-----WSSSDNVNCRPSLEDPFGINSKX 1045

QY 973 VTFS-----ENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQIQCDI--- 1019

Db 1046 MTISKSEVLKRGTIQDCSTC-----GVATITCSLLPSDLSQVNVSL 1088

QY 1020 ---PFTGIQEBP---NATLXGNLSFDWYIKTSHNHLLIYSTAEILLFENDSVFTLLPQCAF 1073

Db 1089 LMKPTF-IRAHFSSLNLTARGELK-----SENSLTJSSN----- 1123

QY 1074 VRSOTETKVEPFEVNPPLPL--IVGSSVGLLILLALITAALYKLGFPFKQYKMMSE 1128

Db 1124 RXRELAIQISKDGLPGRVPLWILLSAPAGLELLMLLILALWKIGFPFKRPLKKOMEK 1180

Search completed: June 7, 2004, 17:13:03
Job time : 13.9719 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:56:30 ; Search time 39.4889 Seconds
(without alignments)
9084.693 Million cell updates/sec

Title: US-09-902-481B-6
Perfect score: 5875
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGGPPGABQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL 25.*
- 1: sp_archea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_plage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_virus.*
 - 16: sp_bacteriaph.*
 - 17: sp_archeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4387	74.7	1151	11 Q9J130	Q9J130 rattus norv
2	3907.5	66.5	1036	11 Q8CA73	Q8CA73 mus musculus
3	3815.5	64.9	920	6 Q8E984	Q8E984 eus scrofa
4	3484	59.3	1169	4 Q8IVAE	Q8IVAE homo sapien
5	3326.5	56.6	1169	11 Q9QX47	Q9QX47 mus musculus
6	3251.5	55.3	1161	11 Q9QVE7	Q9QVE7 rattus norv
7	1529.5	26.0	1161	11 Q9WTV4	Q9WTV4 mus musculus
8	1519	25.9	1160	11 Q9R200	Q9R200 mus musculus
9	1414	24.1	1196	13 Q98TF1	Q98TF1 cyprinus ca
10	1363.5	23.2	1187	13 Q98TF0	Q98TF0 cyprinus ca
11	1350.5	23.0	1086	4 Q9EH81	Q9EH81 homo sapien
12	1270	21.6	927	6 Q9HZV0	Q9HZV0 bos taurus
13	1163.5	19.8	1167	11 Q8B340	Q8B340 rattus norv
14	1119	19.0	1167	11 Q8B341	Q8B341 rattus norv
15	1052.5	17.9	1188	11 Q7TQC3	Q7TQC3 mus musculus
16	1051	17.9	1171	13 Q42094	Q42094 gallus gall

17	1041	17.7	1038	11 Q8BS01	Q8BS01 mus musculus
18	1014	17.3	895	11 Q9WUP8	Q9WUP8 mus sp. itg
19	1013.5	17.3	1160	6 Q8MKF4	Q8MKF4 felis silve
20	865	14.7	348	4 Q8TES5	Q8TES5 homo sapien
21	852	14.5	1332	5 Q9BPQ8	Q9BPQ8 halocynthia
22	807	13.7	205	11 Q83001	Q83001 rattus norv
23	753.5	12.8	780	13 Q8E271	Q8E271 xenopus lae
24	738	12.6	823	4 Q8WY18	Q8WY18 homo sapien
25	686.5	11.7	823	11 Q8CE84	Q8CE84 mus musculus
26	669	11.4	1032	11 Q61989	Q61989 mus musculus
27	643	10.9	1036	11 Q91YD5	Q91YD5 mus musculus
28	640.5	10.9	1033	6 Q9BGU3	Q9BGU3 bos taurus
29	630.5	10.7	257	11 Q8C270	Q8C270 mus musculus
30	627.5	10.7	1474	5 Q8G687	Q8G687 pseudoplusi
31	619.5	10.5	1041	5 Q9UB90	Q9UB90 lytechinus
32	614.5	10.5	1041	5 Q76378	Q76378 lytechinus
33	579.5	9.9	1054	5 Q9U6S1	Q9U6S1 strongyloce
34	579	9.9	1034	13 Q98TT7	Q98TT7 gallus gall
35	555.5	9.5	1053	11 Q80YPS	Q80YPS mus musculus
36	550	9.4	1033	13 Q42598	Q42598 xenopus lae
37	546	9.3	1036	6 Q7TRP8	Q7TRP8 equus cabal
38	534	9.1	1016	13 Q91779	Q91779 xenopus lae
39	530	9.0	974	11 Q924W2	Q924W2 rattus norv
40	529	9.0	1073	11 Q8CC06	Q8CC06 mus musculus
41	529	9.0	1119	5 Q8G888	Q8G888 pseudoplusi
42	526	9.0	1047	6 Q9MZD6	Q9MZD6 bos taurus
43	525.5	8.9	1007	6 Q9GK48	Q9GK48 bos taurus
44	522.5	8.9	1132	11 Q8GZ18	Q8GZ18 mus musculus
45	512.5	8.7	1034	6 Q9TUN4	Q9TUN4 oryctolegus

ALIGNMENTS

RESULT 1

Q9J130 PRELIMINARY; PRT; 1151 AA.

AC Q9J130 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Integrin beta 2 alpha subunit.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RA Fathallah D.M. Sr., Zeria K. Jr.;

RT "Cloning of the rat CD11b cDNA sequence."

RL Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AF268593; AAF81280.1; ;

DR HSSP; P11215; 1BHQ.

DR GO; GO:0008305; C:integrin complex; IEA.

DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.

DR GO; GO:0007160; P:cell-matrix adhesion; IEA.

DR InterPro; IPR000413; Integrin_alpha.

DR InterPro; IPR002035; VWF A.

DR Pfam; PF01839; FG-GAP; 3.

DR Pfam; PF00357; integrin_A; 1.

DR Pfam; PF00092; vwa; 1.

DR	PRINTS; PR01185; INTEGRINA.
DR	PRINTS; PR00453; VWFADOMAIN.
DR	SMART; SM00191; Int_alpha; 5.
DR	SMART; SM00327; VWA; 1.
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR	PROSITE; PS00234; VWF A; 1.
DR	SEQUENCE 1151 AA; 126943 MW; 87856595D047CA5 CRC64;

Query Match 74.7%; Score 4387; DB 11; Length 1151;
Best Local Similarity 73.1%; Pred. No. 3.4e-311;
Matches 831; Conservative 146; Mismatches 158; Indels 2; Gaps 2;

Qy 1 FNLDTENAMTFQENARFGQSVVQLQGSVVVGAPOEIVAAQNGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARFGQSVVQLQGSVVVGAPOEIVAAQNGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPEAVNMSGLSLAATTPPOLLACGPTVHQTCSNTYVKGCLFPGSNLRQPOK 120
Db 77 PLQVPEAVNMSGLSLAATTPPOLLACGPTVHQTCSNTYVKGCLFPGSNLRQPOK 136
Qy 121 PPEALRGCPQSDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQLKXSKTLFSLMOYSEEF 180
Db 137 PPEALRGCPQSDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQLKXSKTLFSLMOYSEEF 196
Qy 181 RIHFTTFEQQNPNRPSIAKPEITOLLGRTHATGVRKVRIRBELNITNGARKNAFKILIVI 240
Db 197 RIHFTTFEQQNPNRPSIAKPEITOLLGRTHATGVRKVRIRBELNITNGARKNAFKILIVI 256
Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGAPEFRSEKSRQELNTIASKEPRDHVQVN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGAPEFRSEKSRQELNTIASKEPRDHVQVN 316
Qy 301 NFEALKTIONLRKPIFAIGTGTGSSSEHMSQGFSAITSGNGLLSTGVSQYDWAG 360
Db 317 NFEALKTIONLRKPIFAIGTGTGSSSEHMSQGFSAITSGNGLLSTGVSQYDWAG 376
Qy 361 GVELYTSKESPFINMTVDSDMDAYLGXAAIILNRVQSLVLAGPRYOHIGLVAWFR 420
Db 377 GVELYTSKESPFINMTVDSDMDAYLGXAAIILNRVQSLVLAGPRYOHIGLVAWFR 436
Qy 421 QNTGWESNANVKGQIGAYFGASLCSVDMSNGSTDVLVIGAPHYBQTRGGQVSVCP 480
Db 437 QNTGWESNANVKGQIGAYFGASLCSVDMSNGSTDVLVIGAPHYBQTRGGQVSVCP 496
Qy 481 PRGQBARQCDVLYGEOQPKRGALTVLGDVNGDKLTDVAIGARBEEDNCGAVLF 540
Db 497 PRGQBARQCDVLYGEOQPKRGALTVLGDVNGDKLTDVAIGARBEEDNCGAVLF 555
Qy 541 HGTSGSGSPSHSQRIAGSKLSPRIQYFGQSLGSGQDITMDGLVLTVGAQCHVLLRSQ 600
Db 556 HGTSGSGSPSHSQRIAGSKLSPRIQYFGQSLGSGQDITMDGLVLTVGAQCHVLLRSQ 615
Qy 601 PVLRVKALMEFNPRVARNVFCNDQVVKGEAGVRVCLVHVKSTRRLBEGQIQSVVT 660
Db 616 PVLRVKALMEFNPRVARNVFCNDQVVKGEAGVRVCLVHVKSTRRLBEGQIQSVVT 675
Qy 661 YDLALDSGRPHRAVFNETKSTRTQCTGLGTCTCLKQLPNCIEDPVSPIVLRNF 720
Db 676 YDLALDSGRPHRAVFNETKSTRTQCTGLGTCTCLKQLPNCIEDPVSPIVLRNF 735
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLFTALPPPEKNCNDNIODDLSITFSWSDCLVVG 780
Db 736 SLVGTPLSAFGLNRPVLAEDAQRLFTALPPPEKNCNDNIODDLSITFSWSDCLVVG 795
Qy 781 GPRENVTVVNDGEDSVRQTVTPFFDLDSYRVKSTLQNSORSNRLACESASSTEV 840
Db 796 GPRENVTVVNDGEDSVRQTVTPFFDLDSYRVKSTLQNSORSNRLACESASSTEV 854
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITDVEDVDSKASLGNKLLKANVTSENNPRTKTEF 900
Db 855 SGALKSTSCSINHPIPPENSEVTFNITDVEDVDSKASLGNKLLKANVTSENNPRTKTEF 914
Qy 901 QLELPKVIAYVWVTSHGVTYKLNFTASENTRVQWQYQVSNLQBSLPSLFLVFPV 960
Db 915 QLELPKVIAYVWVTSHGVTYKLNFTASENTRVQWQYQVSNLQBSLPSLFLVFPV 974
Qy 961 RLKQTVIWRPQVTFSENLSSCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
Db 975 RLKQTVIWRPQVTFSENLSSCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1034
Qy 1021 PFGIQEENATUKGNLSPWYIKTSHNHLIYSTABILNDSVFTLLPQGAQVFSQET 1080
Db 1035 PFGIQEENATUKGNLSPWYIKTSHNHLIYSTABILNDSVFTLLPQGAQVFSQET 1094
Qy 1081 KVEPPEVNPPLIVGSSVGLLILALITAAALYKLGFFKQYKQKQWMSGEGPQAPQ 1137

Db 1095 KVEPPEVNPPLIVGSSVGLLILALITAAALYKLGFFKQYKQKQWMSGEGPQAPQ 1151
RESULT 2
Q8CA73 PRELIMINARY; PRT: 1036 AA.
ID Q8CA73 AC Q8CA73
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Integrin alpha M.
GN ITGAM OR F730045J24RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
DR EMBL; AK039444; BAC30350.1; -
DR PIR; PT0572;
DR PIR; PT0633; PT0697;
DR PIR; PT0697; PT0697;
DR MED; MGI:96607; Itgam.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 1.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWPA; 1.
SQ SEQUENCE 1036 AA; 115126 MW; 4F4F39BF1E188E77 CRC64;
Query Match 66.5%; Score 3907.5; DB 11; Length 1036;
Best Local Similarity 66.4%; Pred. No. 3.2e-276;
Matches 756; Conservative 121; Mismatches 142; Indels 119; Gaps 2;
Qy 1 FNLDTENAMTFQENARFGQSVVQLQGSVVVGAPOEIVAAQNGSLYQCDYSTGSCPEI 60
Db 17 FNLDTEHPMTFQENAKFGQSVVQLGGSVVVAAPOEAKAVNQTGALYQCDYSTSRCHPI 76
Qy 61 RLQVPEAVNMSGLSLAATTPPOLLACGPTVHQTCSNTYVKGCLFPGSNLRQPOK 120
Db 77 PLQVPEAVNMSGLSLAATTPPOLLACGPTVHQTCSNTYVKGCLFPGSNLRQPOK 136
Qy 121 PPEALRGCPQSDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQLKXSKTLFSLMOYSEEF 180
Db 137 PPEALRGCPQSDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQLKXSKTLFSLMOYSEEF 196
Qy 181 RIHFTTFEQQNPNRPSIAKPEITOLLGRTHATGVRKVRIRBELNITNGARKNAFKILIVI 240
Db 197 RIHFTTFEQQNPNRPSIAKPEITOLLGRTHATGVRKVRIRBELNITNGARKNAFKILIVI 256
Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGAPEFRSEKSRQELNTIASKEPRDHVQVN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGAPEFRSEKSRQELNTIASKEPRDHVQVN 316
Qy 301 NFEALKTIONLRKPIFAIGTGTGSSSEHMSQGFSAITSGNGLLSTGVSQYDWAG 360
Db 317 NFEALKTIONLRKPIFAIGTGTGSSSEHMSQGFSAITSGNGLLSTGVSQYDWAG 376

```

Qy 361 GVFLYTSKEKSTFTINTRVDSMDNDAYLGVAALILNRVQSLVGLAPRYOHIGLVAMFR 420
Db 377 GAFLYTSKDKVTFINTTRVDSMDNDAYLGVAALILNRVQSLVGLAPRYOHIGLVAMFR 436
Qy 421 QNTGWMESNANVKTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTGGQSVCP 480
Db 437 ENPTGWEPTHSIKG----- 450
Qy 481 PRGORARWQCDVAVLYGQGPWGRFGAALTVLGVDNGDKLTDVAIGAPGBEDNRGAVILF 540
Db 451 ----- 450
Qy 541 HGTSGSGISPSHSORLAKSLSPLOYFGOSLSGGQDITMDGLVDLTVGAQGHVLLARSQ 600
Db 451 -----SORIIGAHPSPLOVFGQSLSGKDLTMDGLMDLAVGQGHVLLARSQ 498
Qy 601 PVLKVKAIMENPREVARNPEFCNDQVVKKGAGEVRVCLHVQKSTRDLREGIOQSVVT 660
Db 499 PVLKLEATMEFSPKKVARSVFACQOVLLKXNDAGEVRVCLHVQKSTRDLREGIOQSVVT 558
Qy 661 YDLALDGRPHSRVAFNETKNSRTROTQVGLTQTCETIKLQLPNCIEDPVSPVILBNF 720
Db 559 YDLALDGRPHSRVAFNETKNSRTROTQVGLTQTCETIKLQLPNCIEDPVSPVILBNF 618
Qy 721 SLVGTPLSAPGNLRPVLAEADAQRLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 619 TLVSEPLASFGNLRPVLAEADAQRLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVVG 678
Qy 781 GPRFNVTVVRNCGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWEL-ACESASSTE 839
Db 679 GPQDFNMSVTLRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWEL-ACESASSTE 738
Qy 840 VSGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTE 899
Db 739 GHGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTE 798
Qy 900 FQLELPVKYAVYVMTSHGVSKYLNFTASNTSRVMOHQYQVNLGORSIPISLVEVLP 959
Db 799 FQLELPVKYAVYVMTSHGVSKYLNFTASNTSRVMOHQYQVNLGORSIPISLVEVLP 858
Qy 960 VRLNCTVTVWRPOVTPSENLSCTCTKEPLSHSDFLAELRKA PVNCSIAVCORIQCDI 1019
Db 859 VQINNVTVMDHPQVIFSNLSACHTEQSPHNSFRDQRLTPTVINCVAVCRIQCDL 918
Qy 1020 PFGIOEBFNATLKNLGFWDYIKTSHNHLIVSTAELFNDSVFTLLPGQAFVRSQTE 1079
Db 919 PSFNTQRIQVNTLKNLGFWDYIKTSHNHLIVSTAELFNDSVFTLLPGQAFVRSQTE 978
Qy 1080 TKVEPFPVNPPLPLIVSGVGLLALLITAAALYKLGFFKQYKDMSEGGPPGAEPPQ 1137
Db 979 TKVEPFEVENPVPLIVGSSIGLVILLALITAGLYKLGFFKQYKDMSEGGPPGAEPPQ 1036

RESULT 3
Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
ID Q28984;
DF 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE CD11b (Fragment).
GN CD11b.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactylia; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RA SEQUENCE FROM N.A.
RP Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40072; AAB16869.1; -.
DR HSSP; P11215; 1BHQ.

```

```

DR GO:0008305; C:integrin complex; IEA.
DR GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR00413; Integrin_alpha.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS0234; VWF_A; 1.
FT NON_TER 1
FT NON_TER 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350D5DAC CRC64;

Query Match 54.9%; Score 3815.5; DB 6; Length 920;
Best Local Similarity 79.4%; Pred. No. 1.4e-269; Indels 1; Gaps 1;
Matches 731; Conservative 80; Mismatches 109;

Qy 118 PQFPEALRGCPQBDSDIAFLIDSGSIIPHPFRMKPEFVSTVMEQKKSKTLFSLMQYS 177
Db 1 PQFPEALRGCPQBDSDIAFLIDSGSIIPHPFRMKPEFVSTVMEQKKSKTLFSLMQYS 60
Qy 178 EPRRIHPTKPEFONNPNRSLVKPIIQLLGRTHATGVRKIVRELLNITNGAKNAFKIL 237
Db 61 EDFYTHFTFNDPFRNPSKLLVRPIRQLLGRTHATGIRKVVRELFSKSGARENAKIL 120
Qy 238 IVITDGEKFDPLGYEDVPEADREGVIRVVGDAFRSEKSRQSLNTIASKPPRDHPV 297
Db 121 VVITDGEKFDPLGYEDVPEADREGVIRVVGDAFNWSKSEELNTIASKPPGDHPV 180
Qy 298 QVNNFEALKTIONLREKIPIAEGTGTGSSSEHEHMSQEGFSAATNSGPLLSTVGSYD 357
Db 181 QVNNFEAVKTIONLQEKTEPIAEGTGTGSSSEHEHMSQEGFSAATNSGPLLSTVGSYD 240
Qy 358 WAGVGLYTSKEKSTFTINTRVDSMDNDAYLGVAALILNRVQSLVGLAPRYOHIGLV 417
Db 241 WAGAFLLHFPDRIPIINTRVDSMDNDAYLGVAALILNRVQSLVGLAPRYOHIGLV 300
Qy 418 MFRONTGWMESNANVKTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTGGQSV 477
Db 301 MFKQNSGAKENADIKSGQISYFGASLCSVDVDSNGSTDVLIGAPHYYEQTGGQSV 360
Qy 478 CPLPRGORARWQCDVAVLYGQGPWGRFGAALTVLGVDNGDKLTDVAIGAPGBEDNRGAV 537
Db 361 CPLPQG-RAKWQCRVILCGEQHPWSPFGAALTALGVDNGDKLTDVAIGAPGBEDNRGAV 419
Qy 538 YLFHGTSGSGISPSHSORLAKSLSPLOYFGOSLSGGQDITMDGLVDLTVGAQGHVLL 597
Db 420 YLFHGTSELGISPSHSORLAKSLSPLOYFGOSLSGGQDITMDGLVDLTVGAQGHVLL 479
Qy 598 RSQPVLRVKAIMENPREVARNPEFCNDQVVKKGAGEVRVCLHVQKSTRDLREGIOQS 657
Db 480 RSQPVLRVEASVMPKPREVARNPEFCNDQVVKKGAGEVRVCLHVQKSTRDLREGIOQS 539
Qy 658 VVITYDALDGRPHSRVAFNETKNSRTROTQVGLTQTCETIKLQLPNCIEDPVSPVILR 717
Db 540 IITYDALDGRPHSRVAFNETKNSRTROTQVGLTQTCETIKLQLPNCIEDPVSPVILR 599
Qy 718 LNFSLVGTPLSAPGNLRPVLAEADAQRLFTALFPPEKNCNDNICQDDLSITFSFMSLDCL 777
Db 600 LNFSLVGTPLSAPGNLRPVLAEADAQRLFTALFPPEKNCNDNICQDDLSITFSFMSLDCL 659
Qy 778 VVGGPREFNVTVVRNCGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWELACESASS 837
Db 660 VVGGPREFNVTVVRNCGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWELACESASS 719
Qy 838 TEVSGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNK 897
Db 720 TEVSGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNK 779
Qy 898 TEFQLELPVKYAVYVMTSHGVSKYLNFTASNTSRVMOHQYQVNLGORSIPISLVEL 957

```

```

780 TEFQLELPVKAVVWVTSLEVSTKYENFTASEKTRHVIHQCFNNLQKLPISVVFV 839
958 VPVRLNQTIVNDPQVTFSENLSTCKTKRLPSHSDFLAELRKA PVVNCIAVCORIQ 1017
840 VPVRLNRTVWDQCVTFQSNLRSCTEIRGPHSDFLKLTPTVLNCSIAVCQKIQ 899
1018 DIPFGIQEEFNATLKGNLSP 1038
900 DIPSGIQEELKVTLKGNLSP 920

RESULT 4
Q81VA6 PRELIMINARY; PRT; 1169 AA.
AC Q81VA6;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Integrin, alpha X (Antigen CD11C (p150), alpha polypeptide).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC038237; AAH38237.1; -.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; vwa; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
SQ SEQUENCE 1169 AA; 128521 MW; 1A7B484FEFC79EB6 CRC64;

```

```

Query Match
Best local Similarity 59.3%; Score 3484; DB 4; Length 1169;
Matches 694; Conservative 138; Mismatches 291; Indels 6; Gaps 4;

```

```

QY 1 FNLDTENAMTFQENARFGQSVVQLGSRVWVGAPQEI VAAVNRGSLYOCYSTGSCPEI 60
DB 20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVGAPQKITAANQTGGLYQCYSTGACPEI 79
QY 61 RLQVPVEANMSLGLSLAATTSPQLLAGCTVHQCSENTYVKGICFLFGSLNRQQPOK 120
DB 80 GLQVPEANMSLGLSLASTSPQLLAGCTVHHECGRNMTLTGLCLLGT -QLTOR 137
QY 121 FPEALRGCPQEOESDIAFLIDGSGIIPHDPRMKFEVSTVMEQLKSKTLPSLMQVSEF 180
DB 138 LPVSRQECPRQEQDIIVFLIDGSGSISSRNFAVNFVRAVISQFQRPSTQFSLMQFSNKF 197
QY 181 RIHFTKEFQNNPNRSLVKPTOLLGRHTATGVRKVIRELLNITNGARKAKFLIYI 240
DB 198 QTHFTFEFRSSNPISLLASVHQLGQFTYATAIGNVHRUHFPHASYGARRAAKFLIYI 257
QY 241 TDCEKPGDPLGYDVIPEADREGVIRVIGVGDAFRSEKSRQELNITIAKPRDHVQYN 300
DB 258 TDCKKGBSDLYKDVLPFADMAAGIIRYALGVGLAFQNRNSWKLNDIAKSPQSEHIFKVE 317
QY 301 NFAKTKIQNLREKIFAIBGTOTGSSSPHEHMSQEGFSAITSGPLISTVGSYDNAG 360
DB 318 DFDALQIQNLREKIFAIBGTOTGSSSPHEHMSQEGFSAITSGPLISTVGSYDNAG 377

```

```

QY 361 GVFLYTSKEKSTPIKTRVDSMDNAYLGYAAATILNRVQSIATLGPAYQHIGLVAMER 420
DB 378 GAFLYPPNNMSPFIINMQENVDNRSDYLGSTELALWKGVSQSLVLCAPRYQHTGAIVT 437
QY 421 QNTGMWESNARYKGTQIGAYFGASCLSDVDVDSNGSTDVLVLI GAPHYRYOTRGQVSVCP 480
DB 438 QVSRQWRMEAEVTGTQIGSYFGASCLSDVDVDSNGSTDVLVLI GAPHYRYOTRGQVSVCP 497
QY 481 PRQORARWOCDAVLYCEQCPHGRFGAALTULGDVNGDKLTDAVIGAPCEENRGAIVLF 540
DB 498 PRGWR - RWWCDVLYGEGHPWRFGAALTULGDVNGDKLTDAVIGAPCEENRGAIVLF 556
QY 541 HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLSGGQDLTMDGLVELTVGAQGHVLLRSQ 600
DB 557 HGVLPFSISPSHSORLAGSKLSPRLQYFGQSLSGGQDLTMDGLVELTVGAQGHVLLRSQ 616
QY 601 PVLVRKALMEFNPREVARNVPECDQVYKKEAGEVTVCLHVOKSTRDLRREGQISVVT 660
DB 617 PVLWVGVSQFTPAEIPRSAFECREQVSEQTLVQSNICLYIDKRKNLLGSRDLQSSVT 676
QY 661 YDLALDGRPHSRVAFNETKNSRROTQVGLTQTCETLKLQLPNCIEDPVSPIVLELNF 720
DB 677 LDIALDPGRSLPRAATQETKNSRSLRVVGLKHCENFNLLPSCVEDSVTPTILRNF 736
QY 721 SLVGTPLSAFGLNLRPVLAEDAQELPTALPFFKNCNDNICODDLITITPSFMSLDCLVVG 780
DB 737 TLVGKPLLAFLNLRPMLAADAQRYFTASLPFEKNCADHICQDNLGISFPFGLKSLVG 796
QY 781 GPREFVTVTVRNDGDSYRTQVTEFFPLDLSVRKYSTLQNRQSRQSWRLACSSASTEV 840
DB 797 SNLELNVEVMVNDGDSYRTQVTEFFPLDLSVRKYSTLQNRQSRQSWRLACSSASTEV 854
QY 841 SGALKSTSCSINHPPIPPENSEVPTNITFDVDSKASLGNKLLKANVTSENNMPTNKTET 900
DB 855 SQGTWSTSCRINHLIPRGAQITFLATFDVSPRAVLGDRILLTANVSENNTPRTSKTTF 914
QY 901 QLELPVKAVVWVTSLEVSTKYENFTASEKTRHVIHQCFNNLQKLPISVVFV 959
DB 915 QLELPVKAVVTVVSSHEQFTKYLNFSESEKSHVAMHYQVNNLQCORLPSVINFVWP 974
QY 960 VRLNQTIVNDPQVTFSENLSTCKTKRLPSHSDFLAELRKA PVVNCIAVCORIQCDI 1019
DB 975 VELNQBAVMDVEVSHPNQPSLRCSSEKIAPPASDPLAHIKQNPVLDCSIAGCLRFCDV 1034
QY 1020 PFGIQEEFNATLKGNLSPFYIKTSHNHLIIVSTAEILENDSVFTLLPGQAFVRSQTE 1079
DB 1035 PPSVQVEELDTLKGNLSPFGVRQILQKVSIVWSVAEITFTSVISQLPQGEAFWRAQTT 1094
QY 1080 TKVPEEPVNPPLIVGSSVGLLALLALITAAALYKLGFFKQYKDMSE 1128
DB 1095 TVLEKVKVHNPTPLIVGSSIGGLALLALITAVLYKVGFEKQYKEMEE 1143
RESULT 5
Q9QXH4 PRELIMINARY; PRT; 1169 AA.
AC Q9QXH4;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Leukocyte adhesion glycoprotein p150,95 alpha integrin subunit.
GN ITGAX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang X., Gorski K., Tong C., Rattis F.-M., Teeng S.-Y., Pardoll D.,
RA Tsuchiya H.;
RT Isolation of Genes Selectively Expressed by Dendritic Cells.11;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DB EMBL; AF211864; AAF23492.1; -.

```

DR HSP; P11215; 1BHQ.
DR MG; MGI:96609; Itga9.
DR GO: GO:0008305; C: integrin complex; IEA.
DR GO: GO:0004835; F: cell adhesion receptor activity; IEA.
DR GO: GO:0007160; P: cell-matrix adhesion; IEA.
DR GO: GO:0007229; P: integrin-mediated signaling pathway; IEA.
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF01839; FG-GAP; 3.
DR Pfam: PF00092; vwa; 1.
DR PRINTS: PR01185; INTEGRINA.
DR PRINTS: PR00453; VWFADOMAIN.
DR SMART: SM00191; Int_alpha; 4.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE: PS50234; VWFA; 1.
DR Integrin.
KW INTEGRIN.
SQ SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;
Query Match 56.6%; Score 3326.5; DB 11; Length 1169;
Best Local Similarity 57.2%; Pred. No. 1.1e-233;
Matches 652; Conservative 167; Mismatches 302; Indels 19; Gaps 7;
Qy 1 FNLDENATFQENARGGQSVQIQGSRVVGAPQEIIVANORGSLYQCDYSTGSCPEI 60
Db 20 FNLDKELTHFMDGAEFGHVSQVLDSDSWSVVVVGAPKELKATNQIGGHIKGYHTGCEPI 79
Qy 61 RLQVEEAVNMSGLSLAATTPQQLACGPTVHTQTCENTTYKGLCFLFGSNLRQQPQK 120
Db 80 SLQVPPRAVNIISGLSLAATNPNSWLLACGPTVHTCTRENILYLGCLFLSLSSPQK-QN 138
Qy 121 PEARLGGPQSDIAFLIDGSGSIIIPDRMEKFEVSTWQELKSKTLFSLMQVSEEF 180
Db 139 FFAQOECFKQDDIVFLIDGSGSISSTDFEKLDFVAVVMSQQRSTFSLMQFSDYF 198
Qy 181 RIHTFFEQNNPRSLVKEITQLGRTHETATGVKRVIRELLNITGAKNAPKILIVI 240
Db 199 RVHTFNNFISTSPSLSLGSRVQLRGVYTTASAKHIVITELTQSGARQATKVLIVI 258
Qy 241 TDGKFGDPLGYEDVIPAEDREGVIRVIGVDAPRSEKSEQLNTIASKPPRDHVFQVN 300
Db 259 TDGRKQGNLSYDSVIPMAEASIRIYALGVKAPYNEHSKQELKATASMPSEHVSVE 318
Qy 301 NFEALKTIONLREKIFAETGTGSSSPFHEMSQEGFSAITNSGPLLSTVGSYDWAQ 360
Db 319 NFDALKDIENQLEKIFAETETPSSSTFELEMSQEGFSAVTPDGPVLGAVGSPWSG 378
Qy 361 GVFLYTSKEKSTFINMTVDSDMDAYLGYAAAILRNVRQSLVGLGAPRYOHIGLVAMFR 420
Db 379 GAFLYPSNNRPTFINNSQENEDMRDAYLGYSTALAFWKVGHSLILGAPRHQHTGKVIET 438
Qy 421 QNTGHWESNANVKGTOIGAYFGASLCSDVDNSGSTDVLVLGAPHYEOTRGQVSVCP 480
Db 439 QESRHWRPKSEVRGQIGSYFGASLCSDVDNRDGSITDLVLGVPHYEHTEGGQVSVCFM 498
Qy 481 PRGQARQCDVAVLGEQGPWREGAALTVLGVNGDKLTDVAIGAPGEEDNMGAVLYF 540
Db 499 P-GVGRWHGCGTTLHGEQCHPWRFGAALTVLGVNGDGLADVAIGAPGEENRGAVIF 557
Qy 541 HCTSGSGISPSHQRIACKSLSPLOYGCSISGGDGLTMDGLVNDLTVGAGHVLARSQ 600
Db 558 HGASRQDAPSPQISASQPSRIQYFGQSLGGQDGLTRDGLVDLAVGSRVLLKTR 617
Qy 601 PVLRYKATMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRGGQIQSVYT 660
Db 618 PILRVSPVTHFTPAIRISRVFECQQAPEQTLSDATVCLHIEHSPKQL--GDLSRTVT 675
Qy 661 YDLALDSGRPHRAVFNETKSTRQTVLGTQTCETLKLQLENCEIDPSPVIRLNF 720
Db 676 FDLALDHGRSLSTRAIFKETKTRALTRVKTGLGNHCESYKLLLPACVEDSVPTIRLNF 735
Qy 721 SLVGTPLSAFGNLPVLAEDAQRLEFALPFPPEKNCNDNIQQDLSITFFSFMSLDCLVVG 780

Db 736 SLVGVPISSQLQPLAVDDQTYFTASLPFEKGCADHICQDDLSVWVFPDKTLVVG 795
Qy 781 GREFNAVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQ-----NORSQSRW 829
Db 796 SDLELVNDVTVSNDGSDSYGTTVTLFPVGLSFRVABQVFLRKEDQOWQRGQHSLSH 855
Qy 830 LACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITEDVDSKASLGNKLLKANVTSE 889
Db 856 LMCD--STPDRSQGLWSTSCSRHVIIFRGSGQVFLVTFDVSFPAELGKRLLRARVGE 913
Qy 890 NNMPTNKTEFOLELPVKYAVVAVTVSHGVSTKYLNFTASE-NTSRVMQHYQVSNLQOR 948
Db 914 NNVPFGPTFFQLELPVKYAVVAVTMISSHDQFTKYLNFTSEKETSWEHRFQVNNLQOR 973
Qy 949 SDPISLVFLVPLRNQTVIWDPRQVTFSENLSTCTKRLPSHSDFLAELRKAPVWCS 1008
Db 974 DVEFSINFWVPIELKEAVN-TVMVSHQNPQLTCYRNRLKPTQFDLLTHMQSPVLDCS 1032
Qy 1009 IAVCQRIQCDIPFGIOEBFNATLKNLSGFDWYIKTSHNHLIVSTABILFNDSVFTLLP 1068
Db 1033 IADCLHRCIDIPSLGILDELYFILKNLSFGWISQTLQKVVLLSEBITFNTSVVSQLP 1092
Qy 1069 GQAFVRSQETETKVEPFVNPPLPLIVGSSVGLLLALITAAALYKLGPPKROYKDMSE 1128
Db 1093 GQAFILRAQTKTVLEMYKVHNPVPLIVGSSVGLLLALITAILYKAGFFKROYKMLEE 1152
RESULT 6
QSOYE7
ID QSOYE7 PRELIMINARY; PRT; 1161 AA.
AC QSOYE7;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., Vandervlieten M., Kilgannon P.D., Dietsch G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin."
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
RM EMBL: AF021334; AAF21241.1; -
DR HSP; P11215; 1BHQ.
DR GO: GO:0008305; C: integrin complex; IEA.
DR GO: GO:0004895; F: cell adhesion receptor activity; IEA.
DR GO: GO:0007160; P: cell-matrix adhesion; IEA.
DR GO: GO:0007229; P: integrin-mediated signaling pathway; IEA.
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF01839; FG-GAP; 3.
DR Pfam: PF00092; Integrin_A; 1.
DR Pfam: PF00357; vwa; 1.
DR PRINTS: PR01185; INTEGRINA.
DR PRINTS: PR00453; VWFADOMAIN.
DR SMART: SM00191; Int_alpha; 4.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE: PS50234; VWFA; 1.
KW Integrin.
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;
Query Match 55.3%; Score 3251.5; DB 11; Length 1161;
Best Local Similarity 58.0%; Pred. No. 3.4e-228;
Matches 654; Conservative 158; Mismatches 302; Indels 13; Gaps 9;
Qy 2 NLDTENATFQENARGGQSVQIQGSRVVGAPQEIIVANORGSLYQCDYSTGSCPEIR 61

Db 21 NLDVEEPIVREDAAAFQGVVQFGSRLVVGAPLEAVAVNQTGRLYDCAPATGMCQPIV 80
Qy 62 LOVPVEAVNMSLGLSLAAATTSPPOLLACGPTVHQTCSNTYVKGCLFLFGSNLRQPOKPF 121
Db 81 LRSPLAVNMSLGLSLVATNAQLACGPTAQACVKNYAKGSCLLGLSSL-QFIQAV 139
Qy 122 PRALRCGPQEDSDIAFLIDGSGSIIPHDPFRMKBFVSTVMEOLKXKSTLPSLMQYSEFR 181
Db 140 PASMPCPRQEDMDIAFLIDGSGSIINQDPAQMKDFVKALMGFASTSTLPSLMQYSEFR 199
Qy 182 ILETFKBFONNPNRSLVKPIITOLLGRTHATGVRKVIKRELLANITNGARKNAFKILVIT 241
Db 200 THFTPTFEKNLIDPQSLVDPIVOLQGLTITAGIRTVMEELFHSKXGSKSAKILLVIT 259
Qy 242 DGEKFDPLGDEVDIPEADREGVIRYVIGVGDAFSEKSRQBLNTIASKPPRHHVQVNN 301
Db 260 DQKVRDPLEYSDVIPAADKAGIIRYVIGVGDAFSEKSRQBLNTIASKPPRHHVQVNN 319
Qy 302 FEAALTIQOLREKIPALBGTQGSSESSPHEMSQEGPSAATTSNGLLSTVGSYDWDAG 361
Db 320 FEAALSIQOLREKIPALBGTQGSSESSPHEMSQEGPSAATTSNGLLSTVGSYDWDAG 379
Qy 362 VELYTSKEKSTFNTMRVDSMDNDAYLGAIAAILNENRVQSLVGLGAPRYQHIGLVAMFRQ 421
Db 380 AFIYPENTPTFINNSQENVDMEDSYLGYSTAVAFWKGVHSLILGAPRHOHTGKVIFQ 439
Qy 422 NTGWESNANVKTQICAYFGASLCSVDNSGSDTLVLIGAPHYVYQTRGGGVSCPLP 481
Db 440 EARHWRPKSEVRGTQIGSYFGASLCSVDNSGSDTLVLIGAPHYVYQTRGGGVSCPLP 499
Qy 482 RGRARWQCDVLYGQGOQWGRFGAALTVLGVNDGDKLTVAIGAPGEDNRGAVILFH 541
Db 500 -GVRGWQCEATLHGQHPWEGFVALTVLGVNDGDKLTVAIGAPGEDNRGAVILFH 558
Qy 542 GTSGSGISPSHORIAGSLSPRLVFGQSLSGQDLTMDGLVLTIVGAGHVLRLRSP 601
Db 559 GASRLBIMPSPORVGTGSLSLRQVFGQSLSGQDLTMDGLVLTIVGAGHVLRLRSP 618
Qy 602 VLKVKAIMFNPREVARNVPECDVQVVGKAGEVVRVCLHVOKSTDRDLREGQISVVTY 661
Db 619 LLKVELSIRPAPNEVAKVAVQWERTPTVLEAGEAVCLTVHKGSPDLL--GNVQGSVRY 676
Qy 662 DLALDSGRPHSRVAFNETKXSTRQVTLGLTQCTETLKLQIPLNCIEDVPSVIVRLNFS 721
Db 677 DLALDQRLISRAIFDETNCITLGRKTLGLGPHCETVKLLPDKCEDVAVSPILNFS 736
Qy 722 LACTPLSAFNLPLAEDAQRLPTALPFEKNCNDNICODDLSITPSFMSLDCIIVGQ 781
Db 737 LVKDSASP--RNLHPVLAVGSDHITASLFEKNCNDNICODDLSITPSFMSLDCIIVGQ 795
Qy 782 PREFNVTVTRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSMWLACASASTEVS 841
Db 796 SPFLTIVTVVWNEGSDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSMWLACASASTEVS 854
Qy 842 GALKSTSCSINHPIFENSEVTNIPEDVDSKASLGNKLLKANVTSENMMPTNKTEFQ 901
Db 855 --LESSSCSINHPIFREGATITMTITFDVSYKAFGLDRLLRAKASSENKNDPTNKTAFQ 912
Qy 902 LELPKVAVVAVVTVSHGVSTKYLNFASANTSR--VMQHOVQVSNLQSRSLPISLVLVFP 960
Db 913 LELPKVAVVTVSHGVSTKYLNFASANTSR--VMQHOVQVSNLQSRSLPISLVLVFP 972
Qy 961 RLNQTVIWDROPVTFENLSS--TCHTKERLPSHSDLEALERKAPVNCVSIACQRLQCDI 1019
Db 973 LLNGVAVWD---VTLSSPAQGVSCVSKMPQPPDFLTQIRRSVLDSCSIADCLHFRCDI 1029
Qy 1020 PFFGIQSEFNATLKGMLSPDWYIKTSHNHLIIVSTABILLFNDVSVFTLLPQOGAFVRSQTE 1079
Db 1030 PSDIQDELFIKRLNLSFQWVQVQTLQEKVLLSEAEITPDTSVYSLPQOEAFBAQVBE 1089
Qy 1080 TKVEPFPVNPFLPIVGSVGGILLALITAAIYKLGFFKQYKQMM 1126
Db 1090 TTLEEVVYVEPIFLVAGSSVGGILLALITAAIYKLGFFKQYKQMM 1136

RESULT 7

OSWTV4

ID Q9WTV4

AC Q9WTV4

DT 01-NOV-1999

DT 01-NOV-1999

DT 01-JUN-2003

DE Integrin alpha L

OS ITGAL

GN Mus musculus (Mouse)

OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=DA/2J; TISSUE=Spleen;

RA Ma R.Z., Teuscher C.;

RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF065902; AAD25885.1; --

DR HSP; P20701.1; IFA.

DR MGD; MGI:96606; Itgal.

DR GO; GO:0008305; C: integrin complex; IEA.

DR GO; GO:0004895; F: cell adhesion receptor activity; IEA.

DR GO; GO:0007160; P: cell-matrix adhesion; IEA.

DR InterPro; IPR000413; Integrin_alpha.

DR InterPro; IPR002035; VWF_A.

DR Pfam; PF01839; FG-GAP; 3.

DR Pfam; PF00357; Integrin_A; 1.

DR Pfam; PF00092; vwa; 1.

DR PRINTS; PR01185; INTEGRINA.

DR PRINTS; PR00453; VMPADOMAIN.

DR SMART; SM00191; Int_alpha; 5.

DR SMART; SM00327; VWA_1.

DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.

DR PROSITE; PS00234; VWFA; 1.

DR SEQUENCE 1161 AA; 128240 MW; 86B1027B209E431 CRC64;

Query Match

Best Local Similarity

Matches 396; Conservative 220; Mismatches 456; Indels 101; Gaps 37;

Score 1529.5; DB 11; Length 1161;

Pred. No. 2.5e-102;

1 FNLDTENAMTQENNA-RQFGSVVQLQGSRVVVGAPQELVAAQVQSGLYQCDYSGCEP 59

24 YLNDTRPTQSLAQAGRHFGYQVLIQEDG-VVVGAPGE---GDNTGGYHCHTSSEFCQ 79

60 TELQVPEVAVNMSLGLSLAAATTSPPOLLACGPTVHQTCSNTYVKGCLFLFGSNLRQO 119

80 VSLH-GSNETSKYLQXTLATDAKGSLLACDPLGSRCTCDQNTYLSGLCVLPQSLGPM 138

120 KPPEALRCGPQEDSDIAFLIDGSGSIIPHDPFRMKBFVSTVMEOLKXKSTLPSLMQY 179

139 QMRPAYQCMKGKVDLVLFLFDGSGSLDRKDEKILEFMKDVNRKLSNTSYQFAVQFSTD 198

180 FRIHPTFKEP-QNNPNRSLVKPIITOLLGRTHATGVRKVIKRELLANITNGARKNAFK 238

199 CRTEFTFDYQKONKPNPDLVLSGVPMPFLNTFPAINYVAVHVPKESGAPPDVKLV 258

239 VITDGEKFGDPLGDEVDIPEADREG-----VIRYVIGVGDAFSEKSRQBLNTIAS 291

259 IITDG-----EASDKGNISAAHDIYIIICIGKHFVSVQKQKTLHFASBP 304

292 PRDHVQVNNFEAKTIQNLREKIFATEGTQGSSESSPHEMSQEGPSAATTSNGLL 351

305 VEEFKILDTFEKLDLFTDQRIYIATGNTQRLDTFNMELSSSGISADLSKHAVVG 364

352 TVGSDYDAGGVF-LYTSKEKSTFNTMRVDSMDNDAYLGAIAA-IILNRVQSLVGLGAP 409

365 AVGAKWDAGGLDREDIQQATFVQCEPLTSDVRGGYLGTVVAMTSSRSRPLLAGAPR 424

410 YHIGLVAMFR--QNTGWESNANVKTQICAYFGASLCSVDNSGSDTLVLIGAPHY 467

```

DQ 425 YQHVGVLLFOAPRAGRWQTKIEGTQIGSYFGGLCSVDLDDQGEALLIGAPLFF 484
QY 468 EOTRGQSVVCLPRGQARWQDAVLYGEGQPGWGFAGALTVLGVNGKLTVDVAIGA 527
DQ 485 GEQGGGRVFTY---QRRQSLFEMVSELQGDGPGYPLGRFGAITALTDINGDRLTDAVGA 541
QY 528 PGEEDNAGAVLPHGTSGSPSHSORSIAGSKLSPLRQVPGQSLSGQDLTMDGLVDLT 587
DQ 542 PLEB---QGANVIFNGKPG-GLSPQPSQRIQAGVPGIRWFGRIHGVKDLGGDRADVV 598
QY 588 VQAQHVLLASQPVLRKAIMBENPREVARNFECNDQVVKGEAG-EVRVCLHVQKST 646
DQ 599 VGPEGRVVLLSRPVDVWVTELSPEBIPVHEVECSYAREEQKHGKVLKACFRKPLT 658
QY 647 RDLREGOIQSVVTVYDLALDSGRPHSRAVFNETKNSTRQTOVLGLTQCTETKLQLPNC 706
DQ 659 PQ---FQGRLLANLSTYTLQDGRHRSRGLFPDGGSHLSGNTSIPT-DKSCLDFFHFPIC 715
QY 707 IEDPVSPIVLFNLSLV---GTPLSAFGN-LRPVLAEDAQLFTALFPFERNCGNDNICQ 762
DQ 716 IQDLISPINVSLNFSLEEEGTPRDQKRAMQPIILRPSIHTV-TKEIPFERNCGEDKKE 774
QY 763 DDLISITPSFMSLDCLVVGGP-----REPNTVTVRNDGDSYRTQVTFPPDLDSYKV 816
DQ 775 ANLTLSPPARS-----GPIRLASSASLAVETWLSNSEDAYVWRLDLPFRLSPKV 827
QY 817 STLQNSORSWRKLACESASSTESGAL-KSTSCSINHPIPFENSEVTFNITFDVDSKAS 875
DQ 828 EMLQ---PHSRMPVSCEBL---TEGSLITKILKNVSSPIFACQSVSLQVWNTLNS 882
QY 876 LGNKLLKANVTSEN-NMERTNKTFQLELPVKYAVMVVTSHGVS TKYLNFTASNTSR 934
DQ 883 WEDFVELNGTVCHENENSLOEDNSAATHIPVLPVNLTKREQNSITLYISFTPRGPXTQ 942
QY 935 VMOHQYOV---SNLQORSILPILVPLVRLNQTVIWDPO-VTFSENLS---TCHT 985
DQ 943 QVHQVYQRIOPSAYDHNP-TLEALVGP-----WPHSDDPIYTWVSQTDPLVTC 995
QY 986 KE-RLPSSDFLAEILKAPVNVCSIAVCQRIQCDIPFFGIEQEFNATLKNLSFDWYIKT 1044
DQ 996 EDLKRFSZ---AEQCLPGV-----QPRCPV---RREILQVTVGVLSKEIKA 1041
QY 1045 SHNHLIIVSTABILNDVSTFLPQGAFFVSOTETKVEPEVFNPLPIVGVSSVGLLI 1104
DQ 1042 S-STLSLCSLSVSNSSKHFLYGSKA-SEAQVLVKVDLIHSKEMLHVYVLSGIGGLVL 1099
QY 1105 LALITAALYKLGFFKQVQKDM-KSEGPPGAP 1136
DQ 1100 LFLIFLALYKVGFFRANLKERMEADGGVPNGSP 1132

RESULT 8
Q9R200 PRELIMINARY; PRT; 1160 AA.
AC Q9R200;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Integrin alpha L.
GN ITGAI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J, TISSUE=Spleen;
RA Ma R.Z., Teuscher C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065901; AAD25884.1; -.
DR HSSP; P20701; ILPA.
DR MGD; MGI:96606; Itgal.
DR GO; GO:0008305; C:integrin complex; IEA.

```

```

DR GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
DR SEQUENCE 1160 AA; 128127 MW; A33CS31B139F1FAD CRC64;

Query Match 25.9%; Score 1519; DB 11; Length 1160;
Best Local Similarity 33.6%; Pred. No. 1.4e-101; Indels 100; Gaps 36;
Matches 394; Conservative 219; Mismatches 459;

QY 1 FNLDTENAMTFQENA-RGFGQSVVQLQGSRRVVGAPQEIIVAAHQRSGLYQCDYSTGSCBP 59
DQ 24 YNLDTRPTQSFQAQGRHFGYQVLQIEDG-VVVGADGE---GDNTGGLYHCRTSSEFCQP 79
QY 60 IRLQVPEAVNMSLGLSLAATTSPPQLIACGPTVHQCSTENTYVKGLCFLPGSLRQPPQ 119
DQ 80 VSLH-GSNHTSKYLGMTLATDAAGKSLACDPLGSLRTCDQNTYLSGLCYLFPQSLBGM 138
QY 120 KPPEARLGCQEBDSIAFLIDGSGSIIPHFRMKGFEVSTVMBQKSKTFLFSLMQYSEE 179
DQ 139 QNRPAVQECCKGKGVLDVFLFDGSGQSLDRKDFEKLLEPMKDMVRKLSNTSYQFAAVQFSTD 198
QY 180 FRIHFTPEP-QNNPNRSLVKGITQLLGHRTATGVKRVIRBELNITNGARKAFKILI 238
DQ 199 CRTEFTFLVYKONKDPVLLGVSQPMFLTNTFRAINVYVAHVFEKESGARPDATKVLV 258
QY 239 VITDGEKFGDPLGVEDVIPEADREG-----VIRVYVGVGDAPRSEKSRDELNTIAKSP 291
DQ 259 IITDG-----EASDKGNISAADITRYIIIGIKHVSQKQKTLHFASEP 304
QY 292 PRDHVQVNNFPAKTIQONLRKIFAIETGTQSGSSSFHEHMSQEGFSAAITNGPILLS 351
DQ 305 VBEFVKILDTFEKLDLFTLQRRIVAIETGTRNQLTSFNMELSSSGISADLSKGHAVVG 364
QY 352 TVGSYDAGGVF-LYTSKEKSTFINMTRVDSMDNDVILGVAAA-IILNRVQSVIAGAPR 409
DQ 365 AVGADWAGGFLDREDLQGAITFVGQEPFLSDVRGGTGLGTYVAMWTSRSPSELLAAGAPR 424
QY 410 YQHICLVAMFR--QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHY 467
DQ 425 YQHVGVLLFOAPRAGRWQTKIEGTQIGSYFGGLCSVDLDDQGEALLIGAPLFF 484
QY 468 EOTRGQSVVCLPRGQARWQDAVLYGEGQPGWGFAGALTVLGVNGKLTVDVAIGA 527
DQ 485 GEQGGGRVFTY---QRRQSLFEMVSELQGDGPGYPLGRFGAITALTDINGDRLTDAVGA 541
QY 528 PGEEDNAGAVLPHGTSGSPSHSORSIAGSKLSPLRQVPGQSLSGQDLTMDGLVDLT 587
DQ 542 PLEB---QGANVIFNGKPG-GLSPQPSQRIQAGVPGIRWFGRIHGVKDLGGDRANVV 598
QY 588 VQAQHVLLASQPVLRKAIMBENPREVARNFECNDQVVKGEAG-EVRVCLHVQKST 646
DQ 599 VGPEGRVVLLSRPVDVWVTELSPEBIPVHEVECSYAREEQKHGKVLKACFRKPLT 658
QY 647 RDLREGOIQSVVTVYDLALDSGRPHSRAVFNETKNSTRQTOVLGLTQCTETKLQLPNC 706
DQ 659 PQ---FQGRLLANLSTYTLQDGRHRSRGLFPDGGSHLSGNTSIPT-DKSCLDFFHFPIC 715
QY 707 IEDPVSPIVLFNLSLV---GTPLSAFGN-LRPVLAEDAQLFTALFPFERNCGNDNICQ 762
DQ 716 IQDLISPINVSLNFSLEEEGTPRDQKRAMQPIILRPSIHTV-TKEIPFERNCGEDKKE 774
QY 763 DDLISITPSFMSLDCLVVGGP-----REPNTVTVRNDGDSYRTQVTFPPDLDSYKV 816

```

Db 775 ANLTLSPPARS-----GPLKSSASLAVETLNSGCHDAYVVRDLDPFPRGLSPKRV 827
Qy 817 STLQNRQORSRLACASSTESVSGAL-KSTCSINHPPIPESENTEFNITFDVDSKAS 875
Db 828 ENLQ---PHSRMPVSECEL--TEGSLTKTKNKNSSPIKAGQVSLQVFMFLNLS 882
Qy 876 LGNKLKANTYSEN-NMPTNKTEFQLELPKVAVMVVTSHGVSTKYLNFASENTSR 934
Db 883 WEDFVELNGTVHCENSSQLQEDNSAATHIPVLYPVNLTKEQENSTLYISFTPKGPKTQ 942
Qy 935 VMHQYQVSNLQORSLSPLSLVPLVPLNQTWINDRPO-----VTFSENLS--TCHTK 986
Db 943 QVQHYQV-----RIQSAVDNMT-LEALVGVPRHSDLIITYTWSQVDPLVCHSE 996
Qy 987 E-RLPSSHDFLAELKAPVNCSTIAVCORIQCIDIPFFGIGQEEFNATKGNLSFDWYKTS 1045
Db 997 DLKPS-----SEAPCLPGV--QFRCPIVF---RWEILLQVGTWVLSKEIKAS 1041
Qy 1046 HNHLLIVSTAEILLFNDSTVFTLLPGCAFVRSQTEKTVFPEVNPPLVLCSSVGGLLLL 1105
Db 1042 -STLSCLSSLSVSNSSKHFLYGSKA-SEAQVLVKVDLIHEKMLHYVYLSGGGLVIL 1099
Qy 1106 ALITAAALYKGFPRQYKDMW-SEGGPQCAEP 1136
Db 1100 FLIFLALYKGVFFKRLKKEADGGVNGSP 1131

RESULT 9
Q98TF1 PRELIMINARY; PRT; 1196 AA.
AC Q98TF1
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE CD11-1.
DE C1A1.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxId=7962;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Nakao M., Miura C., Fujiki K., Yano T.;
RT "Molecular cloning of a leukocyte integrin from the common carp.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB048536; EMBL39134.1; -.
DR HSSP; P20701; 1LFA.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004190; P:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0008508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001969; Aspartic-type AS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR020335; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; vwa; 1.
DR Pfam; PF00357; integrin_A; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS0234; VWF_A; 1.
SQ SEQUENCE 1196 AA; 132477 MW; 9369C807E7DCA53B CRC64;

Query Match 24.1%; Score 1414; DB 13; Length 1196;
Best Local Similarity 31.3%; Pred. No. 7,2e-94;
Matches 373; Conservative 227; Mismatches 446; Indels 144; Gaps 39;
Qy 1 FNLDTENAMTFQENARG-FGQSVWLQ-GSR--VVVGAPQETVAANQORSLYQCDSYGS 56

Db 32 FNIDTEHLRFNGAPEDFFGYSVYQTEFGNKRQIIVGAPLE---GNSGTGMYCTADLOS 88
Qy 57 CEPIRLOVP-----VEAVNMSLGLSLAATSPPLLACGPTVHOTCSBNTVYVGLCFPLGS 112
Db 89 CQ--RLQRPQGSVRFFGMSAAVSSAALTS-----CSPYFPECDSVNLGVCYQFSS 140
Qy 113 NLRQOPKPFPEALRGCPQEDSDIAFLIDSGSIIPHDFFRMKGFVSTVMQLKKSKTLFS 172
Db 141 SL-QAVSNFTAAVQECCKREVNLVPLFDGSSMKTVDFENKXNFIDIMKLSNLSKFA 199
Qy 173 LMQYSSEFRHPHTKFPONNPNRSVLPKPTQLLGRHTHATGVKVIKRELLN-ITNGARK 231
Db 200 AVQSTSDVTRVTFPNDVQSGSABEKLWKE--THMSLTNTHKADYILNKLNSKLSADS 258
Qy 232 NAYKILIVITDGKFGEDPLGYED--VIPEADREGVIRYVIGVGDAPFRSEKSRDELNTIAS 289
Db 259 KAQKALVITD---GUPSNDDBYNVLKKCKDEQNILRYIIGV-----KVDLIETQLAS 309
Qy 290 KPRDRHVPQVNNPEALKTIONQREKIFATECTGTGSSSFHEXYSQSGSAALTSNGL 349
Db 310 EPRKNNTFYIKDYSGLKGLLDNLQKLIYNEGSDVAGQDRQKELSGSGSVVYQBSVI 369
Qy 350 LSTVGSYDWAAGVFLYTSKEKSTFINNTRVDSMN-DAYLGYAAAIIILNRVOSLVIGAP 408
Db 370 VGSVGSNDKRGALYEVVG--SGSDFKETEIIDPAVNKDSYMGYSTVLGMRGVSLLPSGAP 428
Qy 409 RYOHIGLVAMFRONTGMBSNANVKGTOICAYFGASLCSVDVDSNGSTDVLIGAPHYB 468
Db 429 RAEHTGLVTLFTKQNTWTVTNSINGEQIGSYFGASLSDLDVDSGDSFLLVGAPLFYQ 488
Qy 469 QTRGGQVSVCPFPRGQARQWQCDVAVYGEQ-----GQPWGRFGAALTVLGDVNGDKLTDV 523
Db 489 SQ-----PRTEGLRVYVLSLSEQKYFKTLNVQSQTTCGFAASVASLXDLNGDGLSDV 540
Qy 524 AIGAPGEEDNRGAVLYPHGTSGSGISPSHS--QRIAGSLSPRIQYFQOSLSGGDLTMDG 582
Db 541 AVGAPLE--NEGVVYIYLGDRTHGINFELTPQISVQSVLPGQCGFVSUTGQMDMNDN 598
Qy 583 LVDLTGCAQGHVLLRSOPVLRVKAIMFNPFRVARNVFCNDQVVKGEAGEVRVCLHV 642
Db 599 LTDIVGAGQGVILLKARFVMSYSAQLSPSKLSNYFECPS--NAPNAFLNLTSCFTV 656
Qy 643 QKSTRDLREGQIQSV--VTVYDLALDSGRPHSRAVFNETKNSTR--RQTVQLGUTQTCET 698
Db 657 TERTSS---TGSLEKKLVNLSNLNVDRVGRMSRGFFPDQSSVSRSTLQOQSVLLDSSGSCFN 713
Qy 699 LKQLPNCIEDPVSVILRLNPSILVGTPLSAFGLNLRPLVAEDAQLFTALPPEKNCND 758
Db 714 FSPMLRCVADTVSPLKIRNFS--QTEMLS--GNSVAVLDVHSKTEENVEVFFQRCNNSN 770
Qy 759 NICODDLSITFSFMSLDCLVVGPPREFNVTVTVRNDGEDSYRTQVTPFFPLDLSYKXST 818
Db 771 NSCVADLKLNFSTN-NTLVVENCAHFTVQVSLANPGDSDSYNTSIVLHYPEGISLSKFA 829
Qy 819 LQNRORSRWLACESASSTEVSGALKSTSCSINHPIFENSEVTNITDV---DSKAS 875
Db 830 IKPSRTR-----SSCGDRDSCGATNRTTCSIDLPIVRSQTTQFGLGTFVGMKWDNDS 881
Qy 876 LGNKLKANTYSENMMPTNKTEFQLELPKVAVMVVTSHGV-STKYLNFASENTSR 934
Db 882 NEMEIMITANSNNNGM---SDTEVRESVQVQFAVDLAI SILVAEDSVTYMNFSLDRGPK 938
Qy 935 VMHQYQVSNLQORSLSPLSLVPLVPLVVR-----LN 963
Db 939 PLNITYKVENSGKDLFVSVTLLPQOTPHVLTPTHTFSMEVHHSFISSYHQIIMCLLN 998
Qy 964 QTVIWDROPVTFSENLSSTCHTKERLPSHSDFLAELKAPVNCSTIAVCORIQ-----C 1017
Db 999 KHLFFSPELSAVQVRSSTGWSLREVTC-SQP--DLNKSSAVHPLTADARLQNVKEYS 1055
Qy 1018 DIPFGIQEE--FNATLKGMLSFDMYIKTSHNHLIVSTAILPNDVSVFILLPQCGAFVR 1075

Db	1056	KSYFBRKONVPSISAEINNTSYLNQTS-----SELKYNPH-----	-R 1094
Qy	1076	SQTEKVEPFEVENPLPIIV-GSSVGLLHLLALITAAKYLGPFKQYKD	1124
Db	1095	SQTEVKVE-FVPPSLMLIVCTGAVGGFFFIILITLLKRCGFFKRNRPD	1143
RESULT 10			
Q98TF0	PRELIMINARY; PRT; 1187 AA.		
AC	Q98TF0;		
DT	01-JUN-2001 (TremBLrel. 17, Created)		
DT	01-JUN-2001 (TremBLrel. 17, Last sequence update)		
DT	01-OCT-2003 (TremBLrel. 25, Last annotation update)		
DE	CD11-2.		
GN	C1A2.		
OS	Cyprinus carpio (Common carp).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;		
OC	Cyprinidae; Cyprinus.		
OX	NCBI_TaxID=7962;		
EN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Peritoneal exudate cells;		
RA	Kimura M., Fujiki K., Nakao M.;		
RT	"Molecular cloning of a leukocyte integrin from the common carp.";		
RL	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB048537; BAB39135.1; ..		
DR	HSSP; P20701; 1LFA.		
DR	GO; GO:0008305; C:integrin complex; IEA.		
DR	GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.		
DR	GO; GO:0004895; F:cell adhesion receptor activity; IEA.		
DR	GO; GO:0007160; P:cell-matrix adhesion; IEA.		
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.		
DR	InterPro; IPR001965; Asparticase AS.		
DR	InterPro; IPR000413; Integrin_alpha.		
DR	InterPro; IPR002035; VWF_A.		
DR	Pfam; PF01839; FG-GAP; 3.		
DR	Pfam; PF00357; integrin_A; 1.		
DR	Pfam; PF00092; vwa; 1.		
DR	PRINTS; PR01185; INTEGRINA.		
DR	PRINTS; PR00453; VWFADOMAIN.		
DR	SMART; SM00191; Int_alpha; 5.		
DR	SMART; SM00327; VWA; 1.		
DR	PROSITE; PS00141; ASP_PROTEASE; 1.		
DR	PROSITE; PS50234; VWFA; 1.		
SQ	SEQUENCE 1187 AA; 131778 MW; 85EDC7CA8B6B1C59 CRC64;		
Query Match 23.2%; Score 1363.5; DB 13; Length 1187;			
Best local similarity 31.5%; Pred.No.3.5e-90;			
Matches 375; Conservative 209; Mismatches 451; Indels 157; Gaps 43;			
Qy	1	FNLDTENAMTPOENARG-FGQSVVQLQ-GSR--VVGAPQEIIVAANQSGSLVQCDYSTGS	56
Db	32	FNIDTEHLRFNGTPEDFFGYSVVYQTEFGNRKQIIIVGAPLE--GNSAGEMYSCCTADLQS	88
Qy	57	CEPIRLQVP-----VEAVNMSLGLSLAATPSFPQLLACGPTVHQTCSENTYVKGCLFLGS	112
Db	89	CK-RLQRFSGSVAFPGMSAAVSAALTS-----CSPFAHECDGNSYNGVCYQNS	140
Qy	113	NLRQCPQKPEALRGCPQEDSDIAFLDGSIIIPDPRMKPEFVSTWQELKSKTLFS	172
Db	141	SL-QAVSNPTAAVQECSSKEVNLVFLPDGSSSMKAVEFDMKNFKIDVWKLSNSSIKFA	199
Qy	173	LMQYSEBFRIHFTPKFEFQNNPNSLVKPIITOLLGRTHTATGVKRVIRELIN-ITNGARK	231
Db	200	AVQFSTEIRTFDFNDYQNSAEELKMBE-RHMKSLTNTYKAINTVLKNVLSVSSGADP	258
Qy	232	NAFKILVITDGEKFGDPLGYED--VIPADREGVIRYVIGVDAPFRSEKSRQELNLTIAS	289
Db	259	NAQKALVITD-----GDPSDNDYDNLINICDQNLIRYIIIGV-----KVDLTTLTQLAA	309
Qy	290	KPPRDHVFQNNFPAKTIQNLREKIPAIETGTGSSSSFEHMSQGFSAIISNGPL	349

Db	310	EPKLNNTFYIQEYVNGKGLDNLNKKIYNIEGSKAHRGRDRQKELSQSGFSVYVQBSVI	369
Qy	350	LSTVGSVDWAGGVFLYT---SKEKSTFIWTRVDSMDNDAYLGYAAAIILRNVRVQSEVLG	406
Db	370	VGSVGSNDWRGALYEVWGSCKFRQTEITDPAN---KDSYMGYSTVLGXRGHVSLSFSG	426
Qy	407	APRYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHY	466
Db	427	APRAHTGLVTLFTKNEBWTVMENINGEQIGSVFGASLSDLLDVSDDGDFLLVAGPLF	486
Qy	467	YE-QTRG-QGVSVCPPLPRGQARWQCDAVLYGQ--GQPMGRFGAALTGLVDVNGDKLTD	522
Db	487	YQSQPRABGELYVYTL-----SEQYSOKTLQSTTGRFATSLASLKDNGDGLSD	535
Qy	523	VAIGAPCEHNRGAVYLFHCTSGSIGSPSHS-ORIASGSKLSPLQYFGQSLSGGQDUTMD	581
Db	536	VAVGAPLE--NEGVIYILGDGTGHNPEHAPQRIIPARSVLPGLQOQGVSLSGQMDMND	593
Qy	582	GLVDLTVGAQGHVLLASQVPLRVKAIMENPREVARNVEPCNDQVVKGEAGEVRVCLH	641
Db	594	NLPDIVITGTQGIIVLLNARFVMSAQLSPNMEISLNYFECPGS--NAPNAPLNTSCPT	651
Qy	642	VOKSTRDLREGOIQSV--VTYDLALDSGRPHRAVFNETKNSTR---RQTVLGLTQICE	697
Db	652	VTERTSS---TGSLEKKLVSLNLNDVVGMSRGFPDPMDSSETLQOSVLLDSGSSCS	708
Qy	698	TLKLQCNCEIDPVSPIVLRNFS-----LVGTPLSAFENLRPLVLAEDAQRIFTALPPEK	753
Db	709	NFSIFMLRCVADTVSPKIRMTFSQTMLSGNSLAVL-DIQSRTEEYEVVL-----PQR	761
Qy	754	NCNDNLICQDDLSITFSFMSLDCLVGGPRBEVTVTVRNDGDSYRQTVFFFPPLDISY	813
Db	762	NC-NSNSCVADLKLNFSTN-DTLVVENQAHFTVLSLANPGDDSYNTSIVLHYPEGLSL	819
Qy	814	RKVSTLQORSQSRWLACASASTEVSGALKSTSCSINHPIPFENSEVETNITFDV---	870
Db	820	SKPDAIKPSRTR-----SSCGDRDSGATNRTTCSINLPVRSYGTITQPLGTFRTK	871
Qy	871	DSKASLGNKLLKANKVTSNNMPTNKTBEQLELPVYAVVMVTSHGQ-STKVLNPTAS	929
Db	872	DYDMSDRKEMTITANSNNGNM-----SDMSVRSIPVQFAVELAISVAEDSVTLNLSLE	928
Qy	930	ENTSRAVMQHQVQVSNLQORSIPISLVLVPVRLNQTVWDRPQVPTFSENTSSTCHTKERL	989
Db	929	DRGPKPLNIIYKVNVNGLKGLPVSVTLSPCQ-----THTVTLTPHNFMSQ	974
Qy	990	PSHSDFLAELRKAPVNC-----SIAYCQRIQCDIPFFGQEE	1027
Db	975	EVHESFISSEYH--IIMCLNKLHLPFSPPLSAVQTRTTGRSLWVC-----VSSISTGEI	1026
Qy	1028	FNATLKGNL-----SFDWYIKTSHNHLIVSTAEILFNDVSFTLLPG-QGA	1072
Db	1027	FRSSV--NLMAEAVLQNVKEYESKYSFY-EPRRDHVFNIS-AELNFTSRYNQSTGLKYN	1082
Qy	1073	FVRSQTEKTEPFPFVNPPLPIVGVSSVGLLILLALITAAKYLGQFFKQYKD	1124
Db	1083	PHRSQTEVKVEFVFPFGRMLIVCTGAVGGFFLIILILLKCGFFKRNRPD	1134
RESULT 11			
Q96HB1	PRELIMINARY; PRT; 1086 AA.		
AC	Q96HB1;		
DT	01-DEC-2001 (TremBLrel. 19, Created)		
DT	01-DEC-2001 (TremBLrel. 19, Last sequence update)		
DT	01-JUN-2003 (TremBLrel. 24, Last annotation update)		
DE	Hypothetical protein.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		

RP SEQUENCE FROM N.A.
RC Tissue=Lymph;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL, BCC08777; AAH08777.1; --
DR GO; GO:000305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR SMART; SM00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWEA; 1.
KW Hypothetical protein.
SQ SEQUENCE 1086 AA; 119223 MW; P6F2546B8C632F9 CRC64;

Query Match 23.0%; Score 1350.5; DB 4; Length 1086;
Best Local Similarity 32.2%; Pred. No. 2.7e-89;
Matches 376; Conservative 186; Mismatches 432; Indels 175; Gaps 37;

QY 1 ENLDTENAMTFQ--ENARGFGSVVLOGSVVVGAPQIVAAQNGSLYQCDYSTGSC 58
DB 26 YNLDVGASFPSPRAGRIFGRVQLV-GNGVIVGAPGE---GNSGSLYQCQSGTCHL 81
QY 59 PISLQVPEAVNMISLGLSAAATSPOLLACGETVHTQCTSENYYKGLCFPGSNLRQOP 118
DB 82 PVTLR-GSNVTSKYLGMTLA--TDP----- 103
QY 119 QKPEALRGCPQEDSDIAFLIDGSGSIIIPDFRMKEPFTVMEQLKSKTFLSLMOYSE 178
DB 104 -----TDGS-----ILPAVQFST 117
QY 179 EFRHFTFKFQNNPRSLVKEITQLLGRTHATANGVKVIRELNMITNGAKNAPKILI 238
DB 118 SYKTEPDFSDYVRKDPDALLKHVKMLLTNTFGAINTVATEVFEELGARPDPATKVL 177
QY 239 VIIDGKFGDPLGYEVIPEADREGVIRVYLVGDAPRSEKSFQELNTIASPPRPHVQ 298
DB 178 IITDGR--ATDSNGNDAKD-----IIRYIIGKHQFQKESQETLHKFASKPASEFVKI 230
QY 299 VAMFEALKTQNLKIKFAIEGTOTGSSSFHEMSQEGFSAATISNGELSLTSGSYDM 358
DB 231 LDTFEKLKOLFTELQKKIYIEGTSKQDLTSFNNELSSGISADLSRKHAVVGAVGAKDW 290
QY 359 AGGVF-LYTSKEKSTHINMTRVDSMDNDAYLGAAA-IILNRVQSLVIGAPRYOHIGLV 416
DB 291 ACGFLDLKADLQDDTFIGNPLETPFVRAGYLVYTVTLPSRQKTSLLASGAPRYQHMGV 350
QY 417 AMFR--QNTGMWESNANVKQTQIGAYFGASLCSDVDNSGSTDVLVIGAPHYYEOTRGQ 474
DB 351 LLLQEPQGGHNSQVQTIHTQIGSVFGELOGVDVDQDETSELLIGAPLFYGEQGR 410
QY 475 VSVCLPLRGQRARQCDAV--LVGEQGPQGRFGAALTULGVNVDKLTVAIGARDEED 532
DB 411 VFIV-----QRRQLGPEVSELOQDPCGYPLGRFGEAITALTIDNGDLVDVAVGAPLEB- 464
QY 533 NRGAVLPHGTSGSGISPSHRSIAGSKLSPRLQYFGOSLGGODLTMDGLVLTVAQOG 592
DB 465 -QGVYIFNGHG-GLSFQSPQRIEGVQLVSGIQWFGRSIHGVKOLEGLADVAVGAES 522
QY 593 HVLLRSQPLRVKAIMFNPVARNVFCNDQV--KGKEAGEVRVCLHVQKSTDRLR 651
DB 523 QMIVLSRPVVDVMTLSFSPAETPVHEVECSYSTSNKMEGVNITICFQI-KSLIPQF- 580
QY 652 EGQIQSVTVYDLADSGRPHSAVFNKSTRTQTOVLGLTQTCETLKLQLENCTEDPV 711
DB 581 QGRVLNITVTLQDKGHRTRRRGLFPGGRHRLRNIAVT-TSMSCDTFSFHPFVCVDLI 639

QY 712 SPIVRLNFSI---VGTPLSAPGN-----LRPVLAEQAQRLFTALPFFKCNKGNINICQ 762
DB 640 SPINVSINFSLWEEGTPRQORACKDIPILRPSLHSETWEI-----PFEKNCCKGDKKE 694
QY 763 DDLSTTFMSLDCLVVGGPREFNVTVVRNDGSDSYRTQVTFPFFDLSTLSTKYSTLQNO 822
DB 695 ANLAVSFSPARSRALRLTAPASLSVLSLSNLEBDAYVWQLDHFPPGLSFRKVEML--- 751
QY 823 RSQSRWLACHS--ASSTEYSGALKSTSCSINHPIPPENSSEVTNITPDVDSKASLGNKL 880
DB 752 KPHSQIPVSCBELPEERLLSRAL---SCNVSSIFRAGHSVALQMMFNTLVNWSGDSV 808
QY 881 LKANVTSENN-----MPRTNKTEPQLBLPVKAYVMVVTSHGVSTKYNLFTASNTSRV 936
DB 809 ELHANVTNNEDSLLEDNSATTI---IPILYPINILIQDQEDSTLYVSTFKPGPKHQV 865
QY 937 QHQYQV---SNLQORSPL-ISLVELVPRLNQTWINDRPQVTFPSENLSSTCHTK--BELP 990
DB 866 KMTQVRIQPSIHQHNITPLEAVVGPFPSPSEGPITHQSVQMEPPV--PCHYEDLEKLP 923
QY 991 SHSD--FLAELRKAPVWNCSTIAVCQRIQCDIPFPGIQEENATLKGNIASFQYINTSHH 1049
DB 924 DAARFCLPGALFRCPVV-----FRCEILVQVIGTLELVGEIEAS-SM 964
QY 1049 LLIVSTABILFNDVSFTLLPGQAFVRSOTTKVEPPEVNPPLFLIVGSSVGGLLLLALI 1108
DB 965 FSLASSLSISFSSSKHFLYGSNASL-AQVVMKYDVVYKQMLYLYLVSIGGGLLLLLLI 1023
QY 1109 TAALYKLGFFKQVQKMMSEG-GPPGAP 1136
DB 1024 FIVLYKGVFFRNLKEMAGRGVNGIP 1052

RESULT 12
Q8HZV0 PRELIMINARY; PRT; 927 AA.
AC Q8HZV0;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Lymphocyte function-associated antigen 1 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Thumbikat P., Kannan M.S., Maheswaran S.K.;
RT "Sequence of the alpha subunit of bovine lymphocyte function-
associated antigen 1."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF440778; AAN63636.1; --
DR FIR; A32039; A32039.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR SMART; SM00191; Int_alpha; 4.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWEA; 1.
FT NON_TER 1
FT NON_TER 927
SQ SEQUENCE 927 AA; 102523 MW; 02E2CF09917081EC CRC64;

Query Match 21.6%; Score 1270; DB 6; Length 927;
Best Local Similarity 34.7%; Pred. No. 1.6e-83;
Matches 339; Conservative 167; Mismatches 386; Indels 86; Gaps 28;

Db 489 -EEDAFVRIEIEGOMSGYFSGVLCVPDIDMDGTTDFLLVAAPFYHIRGEGRVYVYQVE 547
Qy 483 GQARWQCDVAVLYGQGGQWGRFGAALTGLVDVNGDKLTDVAIGAP-----GEEDNRGA 536
Db 548 -QDASFSLAHTLSGPHGLTNSRFGFAMAAGVDINDQKTDVAIGAPLSPFGAGDGASVGS 606
Qy 537 VYLPHGTSGSGISPHSQRIAGSKLSPRLQVPGQSLGGQDLTMDGLVDLTGVAQGHVLL 596
Db 607 VYIYNGHSG-GLYDPSQOIRASSVASGLHYFGMSVSGGLDFNGDGLADITVGRSDSAVV 665
Qy 597 LRSQPLVRKALMBENPREVARNVPCNDQVYKGEAGEVRVCLHVQKS---TRDRLEEG 653
Db 666 LRSRPVLDLTVGNTTP-----DALPMWFIGM--DYNLCFVDSVVASBPGLREM 715
Qy 654 QIOSVTVYDGLDSDGRPHSRAVENETKSTROTQVGLGTQC----- 696
Db 716 FLNFTVDVQ-----TKQRRLQCEDSSGQCLKRWGSGSLCEHFWLI 760
Qy 697 EFLKQLQFNCIEDPVSPIVRLNFSVINGTPLSAPGNLR-----PVLAEQAQLFTALP--P 750
Db 761 STEEL-----CEEDCFSNITIKVYE-----POTSGRNDYFNPUL--DHYKEPSAIFQLP 809
Qy 751 FFKNGCNDNICODLSITPFSMSLDBLVGGPFRFNVTVVRNDGEDSYRTQVTFPPFLD 810
Db 810 YEKDCNKVFCIAEIQLTN--ISQBELVVGVTKEVTWMTNISLTNSGDSVTMMALNYPRN 868
Qy 811 LSYRKVSTLQONORSORSWPLACESASSTEVSALZSTCSINHPIEPENSEVTNITFDV 870
Db 869 LQFKKI-----QKPSVPDQDDPKPV---ASVILWNCIKIGHPII--KXSSVNVSVTQQL 918
Qy 871 DSKASLGNKLLKANYTSNNPRTNKTBFQLELPKAVVYVMTSHGVSTKYLNTASE 930
Db 919 ESSVFNRTADITVTSNNSKSLARETR--SLQFRHAFIAVLSR--PSVMYMN--TSQ 971
Qy 931 NTSRVWQHOYVSNLQORSLPSLPLVPLVRLNQTIVWDRPQVTSNLSST-----CHT 985
Db 972 SPSDHKPEPFVHGENLFGAVFQLCVPIKLODF-----QIVRVKNLTQDHTCTQ 1025
Qy 986 KEPLPSHSDFLAELKAPVNGSIAVCORIQDIPFGIOEFNATKGLNSFDWYIKTS 1045
Db 1026 SOEPACGSPVQVHXVHVVCAI-----TSNKENVTVAAEISVG 1065
Qy 1046 HNHLLIVSTA-----EILFNDSVFTLPGQAGFVRSQETKVBP-----EVPNPLPLIV 1095
Db 1066 HTKQLLRDVSLEPILGEISFNKSLYGLNAE-----NERTKITVIFLKEETRSPLPII 1119
Qy 1096 GSSVGLLILALITLALYKLGFPKROYKDMWSE 1128
Db 1120 GSSIGGLVLVLIILFKCGFFPKRTQQLNLE 1152

RESULT 14
O88341 PRELIMINARY; PRT; 1167 AA.
AC O88341;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Integrin alpha E2 (Fragment)
OS Rattus norvegicus (Rat)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brenan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between "
RT intraepithelial lymphocytes and dendritic cells in lymph."
RL Eur. J. Immunol. 27:3070-3079 (1997).
DR EMBL; AF020046; AAC23663.1; -;
DR HSSP; P11215; 1BHQ.

GO: GO:0008305; C:integrin complex; IEA.
GO: GO:0004895; P:cell adhesion receptor activity; IEA.
GO: GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWA; 1.
FT NON_TER 1
SQ SEQUENCE 1167 AA; 128593 MW; E3PED7E18B5CBEC CRC64;
Query Match 19.0%; Score 1119; DB 11; Length 1167;
Best Local Similarity 27.8%; Pred. No. 2.5e-72;
Matches 335; Conservative 204; Mismatches 423; Indels 244; Gaps 37;
Qy 47 LYQCDYSTGCEPIRLQVPVEAVNMSLGLSIAATSPQLLAGCPTVH----- 94
Db 67 LYQC-----ATSIIPDEIGQFVHEILMPKRYQVGT 98
Qy 95 -----QTCSENTYVKGLCFPGSNLRQPKPEALRG----- 127
Db 99 LVRNHNGVLVCIQVSRKPSRLNSBELTACSLTLPNLDLQAQAYFSDLEGVLDLGASVNS 158
Qy 128 -----CPQED-----SDIAFLIDGSGSIIPDFFRMKEFV 157
Db 159 GDYYSKGGSTGEBTKSARLRQAVEEEDERAGTEIAIVLDGSGSTEPSDFOKADFI 218
Qy 158 STVMEQL--KKSKTLSLMQYSSEBFRIHFTPKFQNNPNPSLRVKPITQLLGRTHATGV 215
Db 219 STMEVFKECFECFNALVQGVQIQTFFDLDSRDINASLAKVQSIQVQKEVKTASAM 278
Qy 216 RKVIRELLNITNGARKNAFKLIVITGEEKGDPDIGHEDVIPBEADREGVIRVYVIGVDAP 275
Db 279 QHVLNFIPIPSGRSKKALKVMVVLTDGIDIFRDLNLTTVISSKMQGVVRFALGVGNAP 338
Qy 276 RSEKSRBELNTIASKPPREDHVQVNNFENLKTIONQREKIFAETGTOTGSSSSPHEHMS 335
Db 339 ENNTYRELKLIADPRAAHFTKNTYNSALDGLSKLQORLIHMEGT---VGDITLYQLA 395
Qy 336 QEGPSAATITNGP--LLSTVGSYDAGGVFLY--TSKBSKSTFNMTVRVDSMDNA---YLYG 390
Db 396 QTGFSAQILDKGVLLGTVGAFNWSGGALLYNTONGRGRFLNQT--AKEDFRAAQVSYLYG 454
Qy 391 AAAIILNRVQSLVLAGPYQHIGIVAMFRONTGMESNANVKGTQIGAYFGASILCSVDV 450
Db 455 SVALHKAHGVSIVAGAPRHKLRGAVFLOKEDGEEETPWRIEGEQMGSYFGSVLCPVDI 514
Qy 451 DMSGTDVLVIGAPHYETQTRGGQVSVCLPRGQARWQCDVAVLYEGQGPWGRFGAALT 510
Db 515 NWDGITDFLLVAAPFYHIRGEGRVYVYRV--HEQDAPPSLVTLTSGYPLGTSRFGFAMA 573
Qy 511 VLGDVNGKLTDTVAIGAP-----GEEDNRGANVLFHGTSGSGISPHSQRIAGSKLSPR 564
Db 574 AVGDIQNDQKFTDVAIGAPLFGFAGDGASVGSVYIYNGHSG--GLHASTSQIRASSVALG 632
Qy 565 LQYFGQSLSGGGQDLTMDGLVDLTGVAQGHVLLLSRSQPLVRKALMBENPREVARNVPCN 624
Db 633 LYVFGMSVSGGLDFSGDLDLADITVGRSDSAVVLRSPVVDLTVMSTFTPDALP----- 685
Qy 625 DQVVGKGEAGEVRVCLHVQKS---TRDRLEBQIOQSVTVYDGLDSDGRPHSRAVENETKN 681
Db 686 ---NAFKDKMDVELCFKVDSSAIVPSEPLGRGLNSLMTVDVQ-----TKQ 727
Qy 682 STRRQ-----TQVGLGTQTCETLKLQLPNCIEDPVSPIVRLNAPS 721
Db 728 KQRLQCADRSQCCSLMKWSGSSSLCEHFLISTEEL-----CEDDFCSNITIKVSYE 781

QY 722 LVGTPSAFNLRLPVLAEADQRLTALF--PFRKCGNDNICDDLSITFSFMSLCLV 779
DB 782 F-QTSEERNHPNPL--DHYKEPSAIFQLPYEKDCKNVFCTAEIQUTTAISQOD-LVV 837
QY 780 GGPPEFNVTVVRNDGDSYETQVTFPPPLDLSYRKVSTLQNRQSRWSRLACESASSTB 839
DB 838 GITREVTNNTSLTNSGDSYTNALNYPRNLQPKI-----QKPLSPDQDDPKPV- 890
QY 840 VSGALKSTSCINHPFPENSEVTNITFDVDSKASLGKLLKLLKANVTSENNMPRTKTE 899
DB 891 --ASVLNCKIGHPII-KRSSVNVSTWQLEESIFPNRTADITVTISNKSNEKSLARETH 947
QY 900 FOLELPVKYAVYMWVTSHTGVTSKYLNFTASENTSRVMOHQVQVNLGORSIPISLVLPV 959
DB 948 ---SLOPRHIAVLNR--PSWMYNN--TSOSSDHKEFFPNVGENHFGAVFQLOICVP 1000
QY 960 VRLNQTIVDRPQVTFSENLSST-CHTKERLPSSHDFLAEILKAPVNVCSIAVQRIQCD 1018
DB 1001 IITLADL-----QIIRVKHLTKTOHT-----ECTQSGPTCGSDPVQV--- 1039
QY 1019 IPFGIOEFNATL-----KGNLSFDWYIKTSNHLI--VSTABIL-----VNDSVPTL 1066
DB 1040 -----EEMHSVICAITSKENVTAAEISMGTQKLLRDISLQILGEISFNKSUYEG 1092
QY 1067 LPGAFAFVRSQTEKVEFP-----EVPNELPLIVGSSVGGLLLLALITAAALYKLGFFKQY 1122
DB 1093 LNAE-----NHRKITVIFLKEKPHSLPLIIIGSSIGELLAVVVIALLKCGFFKRY 1146
QY 1123 KDMSE 1128
DB 1147 KQANLE 1152

RESULT 15

QY Q7QC3 PRELIMINARY; PRT; 1189 AA.
AC Q7TQC3;
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE All integrin.
GN ITG11.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1] SEQUENCE FROM N.A.
RA Tiger C.-F., Popova S.N., Velling T., Ekman S., Forsberg B.,
RA Gullberg D.;
RT "a1b1 integrin is important for mesenchymal cell function:
RT elimination of a1b1 leads to dwarfism."
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Johansson M., Popova S.N.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY124460; AAM62130.1; --
KW Integrin.
SQ SEQUENCE 1188 AA; 133067 MW; 69B2B2FD75E4271 CRC64;

Query Match 17.9%; Score 1052.5; DB 11; Length 1188;
Best Local Similarity 28.2%; Pred. No. 1.9e-67;
Matches 349; Conservative 184; Mismatches 497; Indels 207; Gaps 45;
QY 1 FNLDENAMTFQ-ENARGGQSVQV--LQGR-VVVGAPQEIIVAAANORGSLYQCDYSTGS 56
DB 23 FNDTRNPRVIAGPNAAPGIVTQHDISGKWLAVVGAPMETNGHQTGDVVKCPVTOGN 82
QY 57 CEPRL-----QVPVEAVNMSIGLSLAATTSPPOLLAGCGPTVHGTCSNTVTKGLCLF 110
DB 83 CTKLNLGRLVTLNVSERKONMELGLSLATPNPDNSFLACSPLWSHEGSSYTTGMCGRV 142

QY 111 GSNLRQOPKFPALRGCTQEDSDIAFLIDGSSIIIPDFRRMKFVSVTMEQ--LKSK 168
DB 143 NSNFRESKTVP-ALQRC-QTYMDIVIVLGSNSIYP--WVEQHEFLINLLKPYIGPGQ 198
QY 169 TLFSLMOYSEPRIHPTFFKEFORNPNRSLVXPIITQLLG-RHTATGVRKVIARELLMITN 227
DB 199 IQVGIVQYGEDAVHEHLNDYRSKDVVRAASHIEQGGTETRTAFGIEFARSEAFQ--K 256
QY 228 GARKNAPKILIVITDGEKFGDPLGYEDVIPEDREGVIRVIGV-----GDARSEKSRQ 282
DB 257 GGRKGAKKVMIVITDGEHSDSP-DLEKVIHQSEKDNVYAVAVLGYNNRGINPETFLN 315
QY 283 ELANTIASKPRDHVFOVNNPEALKTTONLRKIFAIEGTQTGSSSFHEHMQEGFSAA 342
DB 316 EIKYILASDPDKFFNVNDEAALKDVIDALGDRIFSLGINK-NEISFGLMSQTFSSH 374
QY 343 ITNSGELLSTVGSYDNAGGVFLTSKEK-----STFINTRVDSMDNDVLYGAAAILR 397
DB 375 VVEDGILLGAVGAYDNGAVLKETSAGKVIPIHRESLKEPPELKHVAVLYVTTSVWS 434
QY 398 NRV-QSIVLGAPEYQHOHIGLVAMP-RONTGWESNANVKGTOIGAYFGACSCSDVDVNSGS 455
DB 435 SROGRVTVAGAPRNFHTGKVLFSMNNRSLTHQALRGEOISYFGSEITSVDVNDRV 494
QY 456 TDLVLICAPHYBQTR-GGOVSVCPLPRGORARWQCDVLYGCGQPMWGFAGALTVLGD 514
DB 495 TDLVLCAPWYFSEGRERGRVYVNL---RQNFVYNGTLKDSHSYQNAFSGCIASVOD 551
QY 515 VNGDKLTDAVAGPGEDENEGAVYLHGTSGSGISPSHSORISAGSKLSPLOYFGOSLSG 574
DB 552 LNQDSYNDVWVGAPLEDSHRGALYIFHGFO-TMLKPMORITASELAPGLOHFGCSING 610
QY 575 GQDITMDGLVDLTGAGQGHVLLRSQFVLKVAIMEFNPREVARNVF--ECNDQVVKGKE 632
DB 611 QLOLNEGLVDLAVGALGNVAVLWARPVQINASHLHPEPSKI-NIFHKDC-----KR 661
QY 633 AGEVRVCL-----HVQKSTRDLRSGQISQSVTVYDLALDSGPHSAFVNET 679
DB 662 NGRDATCLAAFLCFPIFLAPHFQATVG-----IRYNATMDERMPRAHLDGEG 711
QY 680 KNS--TRRQTOVLGLTQCTETKLQPCIEDPVSPVILRNFSLVGLVTPLSAFGLRPLVA 738
DB 712 GDQFTNRVALLSGQEHQCORINEHVLD-TADYVKPVAFSVEYSLEDP-----DNGPMLD 764
QY 739 EDAQRLTALFPPEKNGCNDNICDDI-----SITFSF 771
DB 765 NGWPTTLRVSPFWFNGCNEDEHCVPDLVLDARSDLPAMEYQCVLGRPAQDCSSYTLSP 824
QY 772 MSLDCLVVGSPRENTVTVRNDGDSYRTQVTFPPPLDLSYKAVSTLQNRQSRWRLA 831
DB 825 DTTVFIESTRRVAVATELENGENAYSAVLISQSENLOF--ASLIQKDDSDNS--IE 880
QY 832 CESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGK-LULLKANVTSEN 890
DB 881 CVNEER-----RLHKKVCNVSYPPFRAKAVAPRLDPEFSKSVFLHHLQIHLGAGSDSHE 935
QY 891 NMPRTNTEFQLELPVKYAVYMWVTSHTGVTSKYLNFTASNTS-----RVWQHOYOV 942
DB 936 QDSTADDNTALLRFLKYHADVLFR---SSLSHFVKANSLSIESYDGLGPPNCVFKV 992
QY 943 SNLQORSLPISLVPL---VPV-----RL-----NOTVIMDRPQVT 974
DB 993 QNLG--PFFHGVWVKITVPIATGNGNLLMLRDPFTDQGNNTSCNIGNSTEYRSTPT-- 1048
QY 975 FSENLSTCHTKERLPQSHD---FLASLRKAPVNVCSIAVQRIQCDIPPGIGIEEFNAT 1031
DB 1049 -BEDLS---HAPQRNHSNDSVWSIICNLRAP-----SOBTSFY 1083
QY 1032 LKGNLSFDWYIKTSNHLIIVSTAEI---LPNDSVFTLLPQGAFAFVRSQTEKVPPEVP 1088
DB 1084 LVGNLWLTSLKALKYRSLKITVNDALQRFHSPFIF-----REEDFSRQVTFEIS 1133
QY 1089 N-----PLPLIVGSSVGGLLLLALITAAALYKLGFFK 1119

1134 QEDWQPIWIVGSTLGLLLALLVIALWKLGPFX 1170

Search completed: June 7, 2004, 17:16:30
Job time : 44.4889 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: June 7, 2004, 16:54:39 ; Search time 49.4608 Seconds
(without alignments)
6495.175 Million cell updates/sec

Title: US-09-902-481B-6
Perfect score: 5875
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKMMSEGGPPGAEPPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5868	99.9	1153	2 AAW65090	AAW65090 Human Bet
2	5868	99.9	1153	3 AAB07360	AAB07360 Human CDI
3	5868	99.9	1153	5 AAU80252	AAU80252 Human int
4	5868	99.9	1153	5 AEG61469	ABG61469 Human Bet
5	5868	99.9	1153	5 AAO14428	AAO14428 Integrin
6	5868	99.9	1153	7 ADD25615	Add25615 Binding d
7	5858	99.7	1153	2 AAR04136	AAR04136 Alpha sub
8	3464	59.0	1163	2 AAR07120	AAR07120 p150.95 a
9	3450	58.7	1163	3 AAW65091	AAW65091 Human CDI
10	3450	58.7	1163	3 AAB07361	AAB07361 Human CDI
11	3450	58.7	1163	5 AEG61470	ABG61470 Human Bet
12	3448	58.7	1163	6 AAB07406	ABU07406 Protein d
13	3417	58.2	1161	2 AAR78166	AAR78166 Human bet
14	3417	58.2	1161	2 AAW23049	AAW23049 Human bet
15	3417	58.2	1161	2 AAW57491	AAW57491 Human bet
16	3417	58.2	1161	2 AAW5089	AAW5089 Human Bet
17	3417	58.2	1161	2 AAW72825	AAW72825 Human alp
18	3417	58.2	1161	2 AAW73342	AAW73342 Human alp
19	3417	58.2	1161	3 AAB07359	ABG61468 Human Bet
20	3417	58.2	1161	5 AEG61468	AAW23064 Human Bet
21	3401.5	57.9	1161	2 AAW23064	AAW65106 Human Bet
22	3401.5	57.9	1161	2 AAW5106	AAW72837 Human alp
23	3401.5	57.9	1161	2 AAW72837	AAW73343 Human alp
24	3401.5	57.9	1161	2 AAW73343	AAW73343 Human alp
25	3401.5	57.9	1161	3 AAB07376	AAW7376 Human alp

26	3401.5	57.9	1161	5 ABG61485	ABG61485 Human Bet
27	3241.5	55.2	1161	2 AAR78169	AAR78169 Rat alpha
28	3239.5	55.1	1161	2 AAW23062	AAW23062 Rat beta
29	3239.5	55.1	1161	2 AAW60004	AAW60004 Rat alpha
30	3239.5	55.1	1161	2 AAW72824	AAW72824 Rat alpha
31	3239.5	55.1	1161	3 AAB07374	ABG61483 Rat alpha
32	3239.5	55.1	1161	5 AEG61483	ABG61483 Rat beta2
33	3232.5	55.0	1161	2 AAW65104	AAW65104 Rat beta-
34	3232.5	55.0	1161	2 AAW73345	AAW73345 Rat alpha
35	3228	54.9	1161	2 AAW23061	AAW23061 Mouse bet
36	3228	54.9	1161	2 AAW60003	AAW60003 Mouse alp
37	3228	54.9	1161	2 AAW65103	AAW65103 Mouse bet
38	3228	54.9	1161	2 AAW72836	AAW72836 Mouse alp
39	3228	54.9	1161	2 AAW73347	AAW73347 Mouse alp
40	3228	54.9	1161	3 AAB07373	ABG61482 Mouse bet
41	3228	54.9	1161	5 AEG61482	ABG61482 Mouse alp
42	3224	54.9	1161	2 AAR78168	AAR78168 Mouse alp
43	3218.5	54.8	1151	2 AAW23059	AAW23059 Rat beta
44	3218.5	54.8	1151	2 AAW60001	AAW60001 Rat alpha
45	3218.5	54.8	1151	2 AAW65101	AAW65101 Rat beta-

ALIGNMENTS

RESULT 1
AAW65090
ID AAW65090 standard; protein; 1153 AA.

XX
AC AAW65090;
DT 28-SEP-1998 (first entry)
DE Human Beta-integrin CD11b subunit protein.

XX Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
XX type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
XX lung inflammation; acute respiratory distress syndrome; CD11b subunit;
XX rheumatoid arthritis.

XX Homo sapiens.

XX US5728533-A.

XX 17-MAR-1998.

XX 07-JUN-1995; 95US-00485618.

XX 23-DEC-1993; 93US-00173497.

XX 05-AUG-1994; 94US-00286889.

XX 21-DEC-1994; 94US-00362652.

XX (ICOS-) ICOS CORP.

XX Van Der Vieren M, Gallatin WM;

XX WPI; 1998-206565/18.

XX Screening assay for modulators of integrin binding - using immobilised or

XX labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.

XX Example 5; Fig 1A-D; 106pp; English.

XX This sequence represents a human beta-integrin CD11b subunit which is
XX used to describe a method for identifying compounds that modulate the
XX interaction of the beta-integrin alpha-d subunit with a binding partner
XX of alpha-d which involves contacting an alpha-d polypeptide with an alpha
XX -d binding partner, one of which is immobilised and the other of which is
XX labelled, in the presence of a test compound, and determining if the
XX compound affects binding between the alpha-d polypeptide and alpha-d
XX binding partner, where the alpha-d polypeptide is alpha-d or its fragment
XX comprising the cytoplasmic, transmembrane or extracellular domain of
XX alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC diseases such as type-1 diabetes, atherosclerosis, multiple sclerosis,
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC and rheumatoid arthritis
XX
SQ Sequence 1153 AA;

Query Match 99.9%; Score 5968; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY	1	FMLDTEAMTFQENARGFQSQVVLQGGSRVVVGAPQEIIVAAHQSLYQCDYSTGSCPEI	60
DB	17	FMLDTEAMTFQENARGFQSQVVLQGGSRVVVGAPQEIIVAAHQSLYQCDYSTGSCPEI	76
QY	61	RLQVPVEANVMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGICLFGSNLRPOQPK	120
DB	77	RLQVPVEANVMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGICLFGSNLRPOQPK	136
QY	121	FPFALRGCEQSDIAFLIDGSGSIIPDFRMRKEFVSTVMEQLEKSKTLFSLMOYSEEF	180
DB	137	FPFALRGCEQSDIAFLIDGSGSIIPDFRMRKEFVSTVMEQLEKSKTLFSLMOYSEEF	196
QY	181	RIHPTFKFQNNPNSLVKPTIOLLGRTHATGVRKVIRELLINTNGARKNAFKILIVI	240
DB	197	RIHPTFKFQNNPNSLVKPTIOLLGRTHATGVRKVIRELLINTNGARKNAFKILIVI	256
QY	241	TDGEKFGDPLGYEDVLPEDRSGVIRYVIGVDAPRSEKSRQELNTIASKPPRDHVFQV	300
DB	257	TDGEKFGDPLGYEDVLPEDRSGVIRYVIGVDAPRSEKSRQELNTIASKPPRDHVFQV	316
QY	301	NFEALKTIONLREKI FAIEGTQTGSSSFEHMGQSGPSAAITNSGPLLSTVSGYDAG	360
DB	317	NFEALKTIONLREKI FAIEGTQTGSSSFEHMGQSGPSAAITNSGPLLSTVSGYDAG	376
QY	361	GVPLYSKESKSTFINNTRVDSQNDAYLGYAAAILLRNVQSLVGLGAPYQHIGLVAMPR	420
DB	377	GVPLYSKESKSTFINNTRVDSQNDAYLGYAAAILLRNVQSLVGLGAPYQHIGLVAMPR	436
QY	421	QNTGWMESNANVKGTOIGAFGASLCSVDVDSNGSTDVLVLI GAPHYYEOTRGQVSVCP	480
DB	437	QNTGWMESNANVKGTOIGAFGASLCSVDVDSNGSTDVLVLI GAPHYYEOTRGQVSVCP	496
QY	481	PRQORARWQDAVLYGEOQPMGRFGAALTVDLVNVDGKLTDAVAGPGEEDNRGAVILF	540
DB	497	PRQORARWQDAVLYGEOQPMGRFGAALTVDLVNVDGKLTDAVAGPGEEDNRGAVILF	556
QY	541	HGTSGSGISPSHSORLAGSKSLPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVILLRSQ	600
DB	557	HGTSGSGISPSHSORLAGSKSLPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVILLRSQ	616
QY	601	PVLRVKALMEFNPREVARNVFECDQVYKGEAGVRVCLHVOKSTRDLRLRGQTSQSVYI	660
DB	617	PVLRVKALMEFNPREVARNVFECDQVYKGEAGVRVCLHVOKSTRDLRLRGQTSQSVYI	676
QY	661	YDLALDSGRPHSAVENETKNSRRTOVLGELTOTCETLKLQLPNCIEDPVPVILRLNF	720
DB	677	YDLALDSGRPHSAVENETKNSRRTOVLGELTOTCETLKLQLPNCIEDPVPVILRLNF	736
QY	721	SLVGTPLSAFNLRLPVLAEQAQLFTALPPEFKNCGNDNI QDDLSITTFPSMSLDCLVVG	780
DB	737	SLVGTPLSAFNLRLPVLAEQAQLFTALPPEFKNCGNDNI QDDLSITTFPSMSLDCLVVG	796
QY	781	GPREFNVTVTRNDGDSYETQVTFEPFLDLSYRKVSTLQNRQSRWLACESASSTEV	840
DB	797	GPREFNVTVTRNDGDSYETQVTFEPFLDLSYRKVSTLQNRQSRWLACESASSTEV	856
QY	841	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKXANTYSENNPRTNKTFF	900
DB	857	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKXANTYSENNPRTNKTFF	916
QY	901	QLELPVKYATVMVTSHGUSTKYLNFNTASENTSRVMQHQYQNSLQORSIPLSLVLPLV	960
DB	917	QLELPVKYATVMVTSHGUSTKYLNFNTASENTSRVMQHQYQNSLQORSIPLSLVLPLV	976

QY	961	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSDPLAELRKAPVWVNCISIAVCQRIQCDIP	1020
DB	977	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSDPLAELRKAPVWVNCISIAVCQRIQCDIP	1036
QY	1021	FFGIQBEFNATLKGNSLDFWYIKTSHNELLIVSTAEILFNDSVFTLLPGQCAFVRSQTET	1080
DB	1037	FFGIQBEFNATLKGNSLDFWYIKTSHNELLIVSTAEILFNDSVFTLLPGQCAFVRSQTET	1096
QY	1081	KVEPFEVNPDLPLIVGSSVGLLILALITAAALYKLGFPKQYKDMSEGPPGPAEPQ	1137
DB	1097	KVEPFEVNPDLPLIVGSSVGLLILALITAAALYKLGFPKQYKDMSEGPPGPAEPQ	1153

RESULT 2

AA07360
ID AA07360 standard; protein; 1153 AA.

XX AA07360;

XX 17-JAN-2001 (first entry)

XX Human CD11b protein sequence.

Human; macrophage infiltration inhibition; alpha_d integrin;
leukocyte integrin; Leu-CAM; leukointegrin; immune response;
inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
lung inflammation; acute respiratory distress syndrome; Crohn's disease;
rheumatoid arthritis; central nervous system injury; CD11b.

XX Homo sapiens.

OS WO200029446-A1.

XX 25-MAY-2000.

XX 16-NOV-1999; 99WO-US027139.

XX 16-NOV-1998; 98US-00193043.

XX 08-JUL-1999; 99US-00350259.

XX (ICOS-) ICOS CORP.

XX Gallatin NW, Van Der Vieren M;

XX WPI; 2000-387751/33.

Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
macrophage infiltration and reduce inflammation at central nervous system
injury sites.

XX Example 5; Fig 1; 270pp; English.

Integrins are a class of membrane-associated molecules that participate
in cellular adhesion. Integrins are made up of an alpha subunit and a
beta subunit. One class of human integrins are restricted to expression
in white blood cells and have a common beta2 subunit: the leukocyte
integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
have an important role in immune and inflammatory responses. The present
protein sequence is the human integrin alpha subunit CD11b. This sequence
was used in an alignment to identify a novel beta2 integrin alpha
subunit: alpha d (AAA60014 and AA07359). The present sequence has
approximately 60% identity to the protein sequence of alpha d. The
Alpha d gene and protein may be useful in therapy for diseases linked to
aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple
sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
(LAD). In addition, anti-alpha d monoclonal antibodies may be used in the
inhibition of macrophage infiltration at the site of a central nervous
system injury. The monoclonal antibodies can also be used to detect and
diagnose Crohn's disease

SQL	Sequence 1153 AA;	
	Query Match	99.9%; Score 5868; DB 3; Length 1153;
	Best Local Similarity	99.6%; Pred. No. 0;
	Matches 1133; Conservative	3; Mismatches 1; Indels 0; Gaps 0;
QY	1 FNLDTENAMTFOENARGFGQSVVLOGSRVVVVGAPQEIIVAAANQSGSLVQCDYSTGSCBPI	60
DB	17 FNLDTENAMTFOENARGFGQSVVLOGSRVVVVGAPQEIIVAAANQSGSLVQCDYSTGSCBPI	76
QY	61 RLQVPVEAVNMSLGLSLAATTPPOLACGPTVHOTCSNTYVKGCLFGLFSGNLROQPOK	120
DB	77 RLQVPVEAVNMSLGLSLAATTPPOLACGPTVHOTCSNTYVKGCLFGLFSGNLROQPOK	136
QY	121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTLPFLSLAQYSEEF	180
DB	137 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTLPFLSLAQYSEEF	196
QY	191 RIHFTFKZPONNPNRSLVKEPTQLGRTHATGVRKVIPELLNITNGARKNAKILIVI	240
DB	197 RIHFTFKZPONNPNRSLVKEPTQLGRTHATGVRKVIPELLNITNGARKNAKILIVI	256
QY	241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRDHVFQVN	300
DB	257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRDHVFQVN	316
QY	301 NFEALKTIONQLREKIPAIETGQTGSSSSPHEMSQEGFSAITNSGPLLSTVGSYDWAG	360
DB	317 NFEALKTIONQLREKIPAIETGQTGSSSSPHEMSQEGFSAITNSGPLLSTVGSYDWAG	376
QY	361 GVFLYTSKEKSTFIMNTRVDSMDNAYLGYAAAIILNRNVSQSLVGLGAPRYOHIGLIVAMFR	420
DB	377 GVFLYTSKEKSTFIMNTRVDSMDNAYLGYAAAIILNRNVSQSLVGLGAPRYOHIGLIVAMFR	436
QY	421 QNTGWNESNANYKTOIGAYFGASLCSVDVDSNGSTDLVIGAPHYETQTRGGQSVCP	480
DB	437 QNTGWNESNANYKTOIGAYFGASLCSVDVDSNGSTDLVIGAPHYETQTRGGQSVCP	496
QY	481 PRGQARWOCDAVLGEOCPQNGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVLYF	540
DB	497 PRGQARWOCDAVLGEOCPQNGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVLYF	556
QY	541 HGTSGSGISPSHSQRIAGSKLSPRLOYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ	600
DB	557 HGTSGSGISPSHSQRIAGSKLSPRLOYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ	616
QY	601 PVLRYKAIKFNPREVARNVFECNDQVVKGEAGEVRVCLHWKSTRDLREGQIQSVVT	660
DB	617 PVLRYKAIKFNPREVARNVFECNDQVVKGEAGEVRVCLHWKSTRDLREGQIQSVVT	676
QY	661 YDLALDSGRPHSRVFNENKSTRTOTVIGLTQTCETLKLQLPNCIEDPVSPVLRNLF	720
DB	677 YDLALDSGRPHSRVFNENKSTRTOTVIGLTQTCETLKLQLPNCIEDPVSPVLRNLF	736
QY	721 SLVGTPLSAFNGLRPVLAEADQRLFTALPPFEKNGNDNICODDLSITFSNLSCLVWG	780
DB	737 SLVGTPLSAFNGLRPVLAEADQRLFTALPPFEKNGNDNICODDLSITFSNLSCLVWG	796
QY	781 GREFNVTVVRNDGSDSVRTQVTFPEPLDLSYRKVSTLQNKORSQSWELACESASSTEV	840
DB	797 GREFNVTVVRNDGSDSVRTQVTFPEPLDLSYRKVSTLQNKORSQSWELACESASSTEV	856
QY	841 SGALKSTSCSINHPIFPENSEVTPNITPDVDGKASLGNKLLKANVTSNNPNRNTKEF	900
DB	857 SGALKSTSCSINHPIFPENSEVTPNITPDVDGKASLGNKLLKANVTSNNPNRNTKEF	916
QY	901 QLELPVKYAVVWVTSYGVSTKVLNFTASENTSRVWQHOYQVSNLQQRSLPLSLVFLVPEV	960
DB	917 QLELPVKYAVVWVTSYGVSTKVLNFTASENTSRVWQHOYQVSNLQQRSLPLSLVFLVPEV	976
QY	961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORIQCDIP	1020
DB	977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORIQCDIP	1036

QY	1021 FFGIOEFNATLKGNLSPDWYIKTSHNHLIVSTABILFNDVSFTLLPGQAFVRSOTET	1080
DB	1037 FFGIOEFNATLKGNLSPDWYIKTSHNHLIVSTABILFNDVSFTLLPGQAFVRSOTET	1096
QY	1081 KVPEFFVFNPLPIVGVSSVGGJLJLLALITAAALYKLGFFKQYKDMSEGGPPGAEPPQ	1137
DB	1097 KVPEFFVFNPLPIVGVSSVGGJLJLLALITAAALYKLGFFKQYKDMSEGGPPGAEPPQ	1153
RESULT 3		
AAU80252		
ID	AAU80252 standard; protein; 1153 AA.	
XX	AAU80252;	
DT	15-JUL-2002 (first entry)	
XX	Human integrin 1 alpha-M subunit protein.	
XX	Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;	
KW	inflammatory disease; autoimmune disorder; Crohn's disease;	
KW	human immunodeficiency virus; HIV; myocardial infarction;	
KW	Sjorgen's syndrome; rheumatoid arthritis.	
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
FR	Misc-difference 499..500	
FT	/note= "Encoded by GGG CAG AGG"	
XX	WO200218583-A2.	
XX	07-MAR-2002.	
XX	31-AUG-2001; 2001WO-US027227.	
XX	01-SEP-2000; 2000US-0229700P.	
XX	(BLOO-) CENT BLOOD RES INC.	
PI	Springer TA, Shimoaka M, Lu C;	
DR	WPI; 2002-382964/41.	
DR	N-PSDB; ABK50046.	
XX	Modified integrin-I or integrin I-like domain polypeptide useful as an	
PT	immunogen to produce antibodies specific to polypeptide, comprises a	
PT	disulfide bond such that polypeptide is stabilized in a desired	
PT	conformation.	
XX	Disclosure; Page 109-112; 112pp; English.	
XX	This invention relates to a modified integrin-I or integrin I-like domain	
CC	polypeptide comprising at least one disulfide bond so that the domain is	
CC	stabilised in a desired conformation. The polypeptide of the invention	
CC	may have antiinflammatory or immunosuppressive activities. The	
CC	polypeptides of the invention have an open conformation and are useful as	
CC	immunogens to produce antibodies that selectively bind to integrin I-	
CC	domain; and for identifying a modulator of integrin activity, or of	
CC	interaction of an integrin and a cognate ligand. The polypeptide of the	
CC	invention, or antibodies (preferably anti-LFA-1 antibody) is useful for	
CC	treating or preventing an integrin mediated disorder which is an	
CC	inflammatory or autoimmune disorder in a subject and for inhibiting the	
CC	binding of an integrin to a cognate ligand such as Crohn's disease,	
CC	nephritis; human immunodeficiency virus (HIV), myocardial infarction,	
CC	Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic	
CC	composition comprising the peptide of the invention is useful for	
CC	treating an integrin mediated disorder in a subject. The polypeptides	
CC	and/or active or antigenic fragments are useful as reagents for diagnosis	
CC	of integrin-mediated disorders. The present sequence represents the human	
CC	integrin-1 alpha-M protein subunit used to generate the mutant	
CC	polypeptides of the invention	

XX	SQ	Sequence 1153 AA;	Query Match	99.9%; Score 5868; DB 5; Length 1153;	Best Local Similarity	99.6%; Pred. No. 0;	Matches 1133; Conservative	3; Mismatches	1; Indels	0; Gaps	0;
Qy	1	FNLDTENAMTFOENARGFGOSVVOLOGSRVVVGAPQBIIVAANORGSLYQCDISTGSCBPI	60								
Db	17	FNLDTENAMTFOENARGFGOSVVOLOGSRVVVGAPQBIIVAANORGSLYQCDISTGSCBPI	76								
Qy	61	RLOVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCFILFGSNLRQPOQK	120								
Db	77	RLOVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCFILFGSNLRQPOQK	136								
Qy	121	FPBALRGCPQEDSDIAFLIDGSGSIIPHDFPRMKEFVSTVMEOLKXSKTLFSLMQYSEEP	180								
Db	137	FPBALRGCPQEDSDIAFLIDGSGSIIPHDFPRMKEFVSTVMEOLKXSKTLFSLMQYSEEP	196								
Qy	181	RIHFTKPEFQNNPNRSLVKEPIITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI	240								
Db	197	RIHFTKPEFQNNPNRSLVKEPIITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI	256								
Qy	241	TDGEKFGDPLGYEDVIPEADREGVIRVIVGVDAFRSEKSRQELNTIASRPPDRHVFQVN	300								
Db	257	TDGEKFGDPLGYEDVIPEADREGVIRVIVGVDAFRSEKSRQELNTIASRPPDRHVFQVN	316								
Qy	301	NFEALKTIQNLREKIPAIETGTQGSSSSEHEHMSQEGFSAATISNGPLISTVGSYDMAG	360								
Db	317	NFEALKTIQNLREKIPAIETGTQGSSSSEHEHMSQEGFSAATISNGPLISTVGSYDMAG	376								
Qy	361	GVFLYTSKEKSTFINMTRVDSMDMDAIVLYGAALILLENRVQSLVGLAPRYQHIGLVAMFR	420								
Db	377	GVFLYTSKEKSTFINMTRVDSMDMDAIVLYGAALILLENRVQSLVGLAPRYQHIGLVAMFR	436								
Qy	421	QNTGWSNSANVKQTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGGQSVSCPL	480								
Db	437	QNTGWSNSANVKQTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGGQSVSCPL	496								
Qy	481	PRGQRAWQCDAVLYGSGQGPWGFAGALTVLGPNVNGDKLTDVAIGAPGEDNRGAVYLF	540								
Db	497	PRGQRAWQCDAVLYGSGQGPWGFAGALTVLGPNVNGDKLTDVAIGAPGEDNRGAVYLF	556								
Qy	541	HGTSGSGISPSHSQRIAGSKLSPRLQFYQGSLSGGQDLTMDGLVDLTVGQAQGHVLLRSQ	600								
Db	557	HGTSGSGISPSHSQRIAGSKLSPRLQFYQGSLSGGQDLTMDGLVDLTVGQAQGHVLLRSQ	616								
Qy	601	PVLRVKALMEENPREVARNPECDQVYVKGKEGEVAVCLHVQKSTRDRREQQIQSVVT	660								
Db	617	PVLRVKALMEENPREVARNPECDQVYVKGKEGEVAVCLHVQKSTRDRREQQIQSVVT	676								
Qy	661	YDLALDSGRPHSRVAFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVSPVILRLNF	720								
Db	677	YDLALDSGRPHSRVAFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVSPVILRLNF	736								
Qy	721	SLVGTPLSAFQNLRPVLAEADAQRLLFTALFPPEKKNQGNNTICQDDLSITTFPSFMSLDCLVWG	780								
Db	737	SLVGTPLSAFQNLRPVLAEADAQRLLFTALFPPEKKNQGNNTICQDDLSITTFPSFMSLDCLVWG	796								
Qy	781	GPREFNVTVTVRNDEGSYRTQVTFPPPLDLSYKRVSTLQNRQSRQSWRLACBSASSTEV	840								
Db	797	GPREFNVTVTVRNDEGSYRTQVTFPPPLDLSYKRVSTLQNRQSRQSWRLACBSASSTEV	856								
Qy	841	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET	900								
Db	857	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET	916								
Qy	901	QLELPVKYAVTVVTVSHGVSTKYLNFTASNTSRVMQHOYQVSNLQORSPLISLVLAVPV	960								
Db	917	QLELPVKYAVTVVTVSHGVSTKYLNFTASNTSRVMQHOYQVSNLQORSPLISLVLAVPV	976								
Qy	961	RLNCTVIWDRPQVTFSENLSTCHTKERLPESHSDFLAEIRKAPVYVNGSIAVCQRIQCDIP	1020								

Db	977	RLNQTVYWRDPQVTFSENLSSCTCHTERLPSHSDFLAELRKAPVWNGSIAVCQRIQCDIP	103
Qy	1021	FFGIOBEFNATLKGNLSPDWYIKTSHNHLLIVSTASILENDVSFTLLPGOGAFVRSOTET	1080
Db	1037	FFGIOBEFNATLKGNLSPDWYIKTSHNHLLIVSTASILENDSVFTLLPGOGAFVRSOTET	1096
Qy	1081	KVPEFPEVPNPLPIVGVSSVGGILLALLITAAALYKLGFFKQYQKDMSEGGPPGAEPQ	1137
Db	1097	KVPEFPEVPNPLPIVGVSSVGGILLALLITAAALYKLGFFKQYQKDMSEGGPPGAEPQ	1153
RESULT 4			
ABG61469			
ID	ABG61469	standard; protein; 1153 AA.	
XX	AC	ABG61469;	
XX	DT		
XX	DT	27-AUG-2002 (first entry)	
XX	DE	Human Beta2 integrin alphaCD11b subunit.	
XX	XX	Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit; LAD;	
KW	KW	leukocyte adhesion deficiency; inflammatory response; diabetes;	
KW	KW	multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;	
KW	KW	inflammatory bowel disease; Crohn's disease; ulcerative colitis;	
KW	KW	immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;	
KW	KW	intracellular cell adhesion molecule; vascular cell adhesion molecule;	
KW	KW	locomotor recovery; locomotor damage; locomotor impairment;	
KW	KW	autonomic dysfunction; sensory dysfunction; spinal cord injury.	
OS	OS	Homo sapiens.	
XX	XX		
XX	XX	WC0200230980-A2.	
XX	XX		
PD	PD	18-APR-2002.	
XX	XX		
FF	FF	15-OCT-2001; 2001WO-US032059.	
XX	XX		
FR	FR	13-OCT-2000; 2000US-00688307.	
XX	XX	(ICOS-1) ICOS CORP.	
PA	PA		
XX	XX	Gallatin WM, Van Der Vieren M;	
PI	PI		
XX	XX	WPI; 2002-463260/49.	
DR	DR		
XX	XX		
PT	PT	Use of an anti-alpha-d monoclonal antibodies for promoting locomotor	
PT	PT	recovery, inhibiting locomotor damage, limiting locomotor impairment, or	
PT	PT	limiting autonomic and sensory dysfunction following spinal cord injury.	
XX	XX		
PS	PS	Example 5; Page 191-194; 270pp; English.	
XX	XX		
CC	CC	The invention relates to promoting locomotor recovery, inhibiting	
CC	CC	locomotor damage, limiting locomotor impairment, or limiting autonomic	
CC	CC	and sensory dysfunction following spinal cord injury by administering an	
CC	CC	anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a	
CC	CC	spinal cord injury victim. The method also involves the use of a ligand	
CC	CC	selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule,	
CC	CC	vascular cell adhesion molecule). The method is useful for promoting	
CC	CC	locomotor recovery, inhibiting locomotor damage, limiting locomotor	
CC	CC	impairment, or limiting autonomic and sensory dysfunction following	
CC	CC	spinal cord injury. In particular, the spinal cord injury comprises	
CC	CC	compression of the spinal cord. The antibodies are also useful for	
CC	CC	reducing inflammation at the site of a central nervous system injury. The	
CC	CC	specification also details the identification of Beta2 integrin alphaD	
CC	CC	cDNAs and proteins, for use in raising the antibodies. Beta2 integrins	
CC	CC	are implicated in diseases such as LAD (leukocyte adhesion deficiency,	
CC	CC	inflammatory response, diabetes, multiple sclerosis, arthritis, graft	
CC	CC	atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative	
CC	CC	colitis, immune complex alveolitis and leukaemia. The present sequence is	
CC	CC	a Beta2 integrin alpha subunit sequence included for comparison with the	
XX	XX	Beta2 integrin alphaD protein sequences	

SQ	Sequence 1153 AA;	
	Query Match 99.9%; Score 5868; DB 5; Length 1153;	
	Best Local Similarity 99.6%; Pred. No. 0;	
	Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;	
QY	1 ENLDTENANTFENARGFGOSVVOLOGSRVVVGAPQEIIVAAHQRSGLYQCDYSTGSCRP 60	
DB	17 ENLDTENANTFENARGFGOSVVOLOGSRVVVGAPQEIIVAAHQRSGLYQCDYSTGSCRP 76	
QY	61 RLQVPVEAVNMSLGLSLAATTSPQLACGPTVHQTCSNTYVYKGLCLFLFGSNLRQOPK 120	
DB	77 RLQVPVEAVNMSLGLSLAATTSPQLACGPTVHQTCSNTYVYKGLCLFLFGSNLRQOPK 136	
QY	121 FPEALRGCTOESDIAFLIDGSGIIPHPFRMKPEFVSTVMEQLKSKTLESIMQYSEEF 180	
DB	137 FPEALRGCTOESDIAFLIDGSGIIPHPFRMKPEFVSTVMEQLKSKTLESIMQYSEEF 196	
QY	181 RHFTPEKEFQNNENPNSLVKPTIQLLGRTHRTATGVKRVIRELINITNGARKNAFKILVI 240	
DB	197 RHFTPEKEFQNNENPNSLVKPTIQLLGRTHRTATGVKRVIRELINITNGARKNAFKILVI 256	
QY	241 TDEKEFGDPLGYEDVTPADREGVIRYVIGVDAPFSEKSRQELNTIASKPPRDHVQFN 300	
DB	257 TDEKEFGDPLGYEDVTPADREGVIRYVIGVDAPFSEKSRQELNTIASKPPRDHVQFN 316	
QY	301 NFELATIQNLREKIFAIEGTOTGSSSPEHEMSQEGFSAATTSNGPLLLSTVGSYDWAG 360	
DB	317 NFELATIQNLREKIFAIEGTOTGSSSPEHEMSQEGFSAATTSNGPLLLSTVGSYDWAG 376	
QY	361 GVFLYTSKEKSTFINTRVDSMDNDAYLGYAAAILLRNVQSVLVGAPRYQHIGLVAMER 420	
DB	377 GVFLYTSKEKSTFINTRVDSMDNDAYLGYAAAILLRNVQSVLVGAPRYQHIGLVAMER 436	
QY	421 QNTGMHESNANVKTQIGAFGLSCVDVDSNGSTDLVLIGAPHYTYQTRGGQVSCPL 480	
DB	437 QNTGMHESNANVKTQIGAFGLSCVDVDSNGSTDLVLIGAPHYTYQTRGGQVSCPL 496	
QY	481 PRQQRARWQCDVLYGEGQGFWRFGAALFVLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540	
DB	497 PRQQRARWQCDVLYGEGQGFWRFGAALFVLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556	
QY	541 HGTSGGSISSHQRIAGSKLSPQLQYFGQSLGGQDLTMDGLVDLTGAGQHVLLRSQ 600	
DB	557 HGTSGGSISSHQRIAGSKLSPQLQYFGQSLGGQDLTMDGLVDLTGAGQHVLLRSQ 616	
QY	601 PVLRVKALMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660	
DB	617 PVLRVKALMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676	
QY	661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQCTETKLQLENCEIDPVSVILRLNF 720	
DB	677 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQCTETKLQLENCEIDPVSVILRLNF 736	
QY	721 SLVGTPLSAFNLRPVLAEDAQRLEFALFFPEKNCNDNICQDDLSITFFSMISLDCILVVG 780	
DB	737 SLVGTPLSAFNLRPVLAEDAQRLEFALFFPEKNCNDNICQDDLSITFFSMISLDCILVVG 796	
QY	781 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSKVSTLQNRQSRQSRWRLACSSASSTEV 840	
DB	797 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSKVSTLQNRQSRQSRWRLACSSASSTEV 856	
QY	841 SGALKSTSCSINHPIFENSNTVENITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900	
DB	857 SGALKSTSCSINHPIFENSNTVENITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916	
QY	901 QLELPVKYAVVMVYTSHGVSSTKYLNFTASENTSRVMQHYQVSNLQORSIPISLVFLVPV 960	
DB	917 QLELPVKYAVVMVYTSHGVSSTKYLNFTASENTSRVMQHYQVSNLQORSIPISLVFLVPV 976	
QY	961 RLNOTVIWDRQVTFSENLSTCTTKERLPSHSDFLAELRKA PVVNCISVACQRIQCDIP 1020	
DB	977 RLNOTVIWDRQVTFSENLSTCTTKERLPSHSDFLAELRKA PVVNCISVACQRIQCDIP 1036	

QY	1021 FFGIOEFNATLKGNLSPDWVIKTSNHELLIVSTABILFNDSVFTLLPFGQAFVRSQTE 1080
DB	1037 FFGIOEFNATLKGNLSPDWVIKTSNHELLIVSTABILFNDSVFTLLPFGQAFVRSQTE 1096
QY	1081 KVEPPEVENPLPLIVGSSVGGILLALITAAALYKLGFFKROYKDMSSGGPPGABPO 1137
DB	1097 KVEPPEVENPLPLIVGSSVGGILLALITAAALYKLGFFKROYKDMSSGGPPGABPO 1153

RESULT 5

AA014428
ID AA014428 standard; protein; 1153 AA.

XX AA014428;

AC AC
XX 03-MAY-2002 (first entry)

XX Integrin Mac-1 alpha subunit.

DE Integrin alpha subunit; variant integrin inserted domain protein;
XX Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
KW open conformation; integrin related inflammatory disorder;
KW integrin related immunological disorder; rheumatoid arthritis; ischaemia;
KW reperfusion; hypovolemic shock; infarction; cerebral shock;
KW viral infection; cancer; gene therapy; vaccine;
KW bioactive agent screening.

XX Unidentified.

OS WO200204521-A2.

XX 17-JAN-2002.

XX 09-JUL-2001; 2001MO-US021805.

XX 07-JUL-2000; 2000US-0216600P.

XX (CALX) CALIFORNIA INST OF TECHNOLOGY.

PA (BLOO-) CENT BLOOD RES.

PI Springer T;

XX WPI; 2002-148167/19.

DR New integrin I domain protein having alteration in at least 2
XX noncontiguous regions and exits in an open conformation, useful for
PT treating, preventing or suppressing inflammatory or immunological
PT disorders.

XX Example 1; Fig 1F; 90pp; English.

XX The invention comprises structurally biased variant integrin inserted (1)
CC domain proteins, wherein the alterations to the protein occur in at least
CC two noncontiguous regions. Specifically the the variant integrin I domain
CC proteins are structurally biased to exist in the open conformation, also
CC thereby altering the binding ability of the protein. The invention also
CC comprises nucleic acids encoding the variant integrin I domain proteins.
CC The integrin I domain proteins and nucleic acids are useful for treating,
CC preventing or suppressing integrin related inflammatory and immunological
CC disorders (e.g. rheumatoid arthritis). The variant integrin I domain
CC proteins and nucleic acids can also be used for treating: ischaemia/
CC reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral
CC infection; and cancer. The variant integrin I domain nucleic acids and
CC proteins may be used in gene therapy, as vaccines and to screen for
CC bioactive agents. The present amino acid sequence represents the Mac-1
CC alpha subunit of integrin

XX Sequence 1153 AA;

Query Match 99.9%; Score 5868; DB 5; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLGGSRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLGGSRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTYVKGCLFPGSNLRQQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTYVKGCLFPGSNLRQQPOK 136
QY 121 FPEARLGCPOEDSIAFLIDGSGSIIPHDPRMKPEFVSTVMEQLKSKTLFSLMOYSEF 180
Db 137 FPEARLGCPOEDSIAFLIDGSGSIIPHDPRMKPEFVSTVMEQLKSKTLFSLMOYSEF 196
QY 181 RIHFTPFKEPQNNPNRSLVKPITQLLGRTHRTATGVRKVIKRELLNITNGARKNAFKILIVI 240
Db 197 RIHFTPFKEPQNNPNRSLVKPITQLLGRTHRTATGVRKVIKRELLNITNGARKNAFKILIVI 256
QY 241 TQGEKPGDPLGYEDVITPEADREGVIRYVIGVGDAPRSEKSRORANTLASKPRDHVFOVN 300
Db 257 TQGEKPGDPLGYEDVITPEADREGVIRYVIGVGDAPRSEKSRORANTLASKPRDHVFOVN 316
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAITSNGLPLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAITSNGLPLSTVGSYDWAG 376
QY 361 GVPLTYSKSKSTPINNTRVDSNDMDAYLGYAAAILRNVRQSLVLCAPRYOHIGLVAMFR 420
Db 377 GVPLTYSKSKSTPINNTRVDSNDMDAYLGYAAAILRNVRQSLVLCAPRYOHIGLVAMFR 436
QY 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYETRGQVSVCP 480
Db 437 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYETRGQVSVCP 496
QY 481 PRQARWQCDVLYGEOGPWGRFGAALTVLGDVNGDKLTVAIGAPGEENRGAVLYF 540
Db 497 PRQARWQCDVLYGEOGPWGRFGAALTVLGDVNGDKLTVAIGAPGEENRGAVLYF 556
QY 541 HGTSGSGISPSHSORTAGSKLSPRLQYFGOSLGGQDLTWDLVLTGCAQGVHLLRSQ 600
Db 557 HGTSGSGISPSHSORTAGSKLSPRLQYFGOSLGGQDLTWDLVLTGCAQGVHLLRSQ 616
QY 601 PVLRVKAIMFNPREVARNVFECDQVYKGEAGVRVCLVHVKSTRDLREGQIOSVT 660
Db 617 PVLRVKAIMFNPREVARNVFECDQVYKGEAGVRVCLVHVKSTRDLREGQIOSVT 676
QY 661 YDLALDSGRPHSAVENETKNSRTRQTVGLTQTCETLKLQLPNCIEDPVSPIVLRANP 720
Db 677 YDLALDSGRPHSAVENETKNSRTRQTVGLTQTCETLKLQLPNCIEDPVSPIVLRANP 736
QY 721 SLVGTPLSAFENLRPVLAEDAQLFTALPFERKNCNDNQCDDLSITTFMSLDCLVWG 780
Db 737 SLVGTPLSAFENLRPVLAEDAQLFTALPFERKNCNDNQCDDLSITTFMSLDCLVWG 796
QY 781 GPREFNVTVRNDGEDSYRTQVTFPPPLDLVSRKYSTLQNSQSRWLACESASSTEV 840
Db 797 GPREFNVTVRNDGEDSYRTQVTFPPPLDLVSRKYSTLQNSQSRWLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKIYAVMVTSHGVSTKYLNFTASENTSRVMQHOVQVSNLQORSPLISLVLPV 960
Db 917 QLELPVKIYAVMVTSHGVSTKYLNFTASENTSRVMQHOVQVSNLQORSPLISLVLPV 976
QY 961 RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVNCSIAVQRIQCDIP 1020
Db 977 RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVNCSIAVQRIQCDIP 1036
QY 1021 PFGIOEBFNATLKNLGFDMYIKTSHNELLIVSTAEILFNDSVTLLPGQAFVRSQTE 1080
Db 1037 PFGIOEBFNATLKNLGFDMYIKTSHNELLIVSTAEILFNDSVTLLPGQAFVRSQTE 1096
QY 1081 KYEPFEPVNPFLPLIVGSSVGGLLALLIITAAALYKLGFFKQYKDMWSEGGPPGAEPQ 1137

Db 1097 KYEPFEPVNPFLPLIVGSSVGGLLALLIITAAALYKLGFFKQYKDMWSEGGPPGAEPQ 1153
RESULT 6
ADD25615
ID ADD25615 standard; protein; 1153 AA.
XX
AC ADD25615;
XX
DT 15-JAN-2004 (first entry)
XX
DE Binding domain-immunoglobulin fusion protein-associated protein #85.
XX
KW Binding domain; immunoglobulin; fusion protein; cytostatic;
antiarthritic; immunosuppressive; antidiabetic; antichryoid;
neuroprotective; hinge region; immunoglobulin heavy chain;
CH2 constant region; CH3 constant region; IgG1;
antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;
rheumatoid arthritis; myasthenia gravis; Grave's disease;
type I diabetes mellitus; multiple sclerosis; autoimmune disease.
XX
OS Unidentified.
XX
PN US2003118592-A1.
XX
PD 26-JUN-2003.
XX
PF 25-JUL-2002; 2002US-00207655.
XX
PR 17-JAN-2001; 2001US-0367358P.
PR 17-JAN-2002; 2002US-00053530.
PR 03-JUN-2002; 2002US-0385691P.
XX
PA (GENE-) GENE-CRAFT INC.
XX
PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
XX
XX WPI; 2003-801317/75.
XX
PT New binding domain-immunoglobulin fusion protein, useful for treating a
subject having or suspected of having a malignant condition or a B-cell
disorder, e.g. melanoma, Grave's disease or autoimmune disease.
XX
PS Disclosure; SEQ ID NO 176; 157pp; English.
XX
CC The invention relates to a binding domain-immunoglobulin fusion protein
comprising a binding domain polypeptide that is fused to an
immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain
CH2 constant region polypeptide that is fused to the hinge region
polypeptide, and an immunoglobulin heavy chain CH3 constant region
polypeptide that is fused to the CH2 constant region polypeptide. The
hinge region polypeptide comprises a wild-type human IgG1 immunoglobulin
hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge
region polypeptide, derived from (a) having 3 or more cysteine residues;
where the mutated human IgG1 immunoglobulin hinge region polypeptide
contains 2 cysteine residues, where the first cysteine is not mutated; a
mutated human IgG1 immunoglobulin hinge region polypeptide, derived from
(a) having 3 or more cysteine residues, where the mutated human IgG1
immunoglobulin hinge region polypeptide contains no more than one
cysteine residue; and a mutated human IgG1 immunoglobulin hinge region
polypeptide, derived from (a) having 3 or more cysteine residues; where
the mutated human IgG1 immunoglobulin hinge region polypeptide contains
no cysteine residues. The binding domain-immunoglobulin fusion protein is
capable of at least one immunological activity comprising antibody
dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The
binding domain polypeptide is capable of specifically binding to an
antigen. Also included are an isolated polynucleotide encoding the
binding domain-immunoglobulin fusion protein, a recombinant expression
construct comprising the polynucleotide (operably linked to a promoter),
a host cell transformed or transfected with a recombinant expression
construct, producing the binding domain-immunoglobulin fusion protein, a

CC	pharmaceutical composition comprising the binding domain-immunoglobulin
CC	fusion protein or polynucleotide and a carrier, and treating a subject
CC	having or suspected of having a malignant condition or a B-cell disorder.
CC	The binding domain-immunoglobulin fusion protein is useful for treating a
CC	subject having or suspected of having a malignant condition or a B-cell
CC	disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,
CC	myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple
CC	sclerosis or autoimmune disease. The present sequence is a binding domain
CC	-immunoglobulin fusion protein-associated protein sequence. Note: The
CC	sequence data for this patent formed part of the printed specification
CC	and is also available in electronic format directly from USPTO at
CC	seqdata.uspto.gov/sequence.html?DocID=2003011892. The authors have not
CC	identified the sequences in the printed specification by their SEQ ID
CC	number therefore none of the sequences can be explicitly identified.
XX	
SQ	Sequence 1153 AA;
Query Match 99.9%; Score 5868; DB 7; Length 1153;	
Best Local Similarity 99.6%; Pred. No. 0;	
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;	
QY	1 FNLDTENAMTFOENARGFGQSVVQLGSRVVVGAPQEIIVAAHQGSLYQCDYSTGSCBPI 60
DB	17 FNLDTENAMTFOENARGFGQSVVQLGSRVVVGAPQEIIVAAHQGSLYQCDYSTGSCBPI 76
QY	61 RLQVPVAVNWSLGLSLAATTSPOLLACGPTVHOTCSNTVVKGLCLFLGSLNRQOQPK 120
DB	77 RLQVPVAVNWSLGLSLAATTSPOLLACGPTVHOTCSNTVVKGLCLFLGSLNRQOQPK 136
QY	121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKBFVSTVMEQLKSKTLPFLMYSBEF 180
DB	137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKBFVSTVMEQLKSKTLPFLMYSBEF 196
QY	181 RIHFTFKFQNNPNRSLVKPTQLLGRTHRTATGVRKVIRELLNITNGARKQAFKILIVI 240
DB	197 RIHFTFKFQNNPNRSLVKPTQLLGRTHRTATGVRKVIRELLNITNGARKQAFKILWI 256
QY	241 TDGEKFGDPLGYEDNTPADREGVIRVYGVGDGAFSEKSRQBLNTIASKPRDHVQVN 300
DB	257 TDGEKFGDPLGYEDNTPADREGVIRVYGVGDGAFSEKSRQBLNTIASKPRDHVQVN 316
QY	301 NFEALXTIQNLRKFIKFAIEGTQTGSSSFEHMSQEGFSAATSNGLLSTVGSYDMAG 360
DB	317 NFEALXTIQNLRKFIKFAIEGTQTGSSSFEHMSQEGFSAATSNGLLSTVGSYDMAG 376
QY	361 GVPLYTSKEKSTPINTRVDSMDNDAYLGAAAIILRNVRQSLVLCAPYOHIGLVAMPR 420
DB	377 GVPLYTSKEKSTPINTRVDSMDNDAYLGAAAIILRNVRQSLVLCAPYOHIGLVAMPR 436
QY	421 QNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYIEQTRGQGVSCPL 480
DB	437 QNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYIEQTRGQGVSCPL 496
QY	481 PRQARWQCDVLYGEGQGFQGRFGAALTVLGDVNGDKLTVDAIGAPGEDNRGAVYLF 540
DB	497 PRQARWQCDVLYGEGQGFQGRFGAALTVLGDVNGDKLTVDAIGAPGEDNRGAVYLF 556
QY	541 HGTSGSGISPSHSORITAGSKLSPLOYFGOSLSGGDLTMDGLVDTLVGAQGHVLLRSQ 600
DB	557 HGTSGSGISPSHSORITAGSKLSPLOYFGOSLSGGDLTMDGLVDTLVGAQGHVLLRSQ 616
QY	601 PVLVRVKAIMEFNPFEVARNVPECNDQVYVKGKAGEVRVCLHVQKSTRDLRREGQISVVT 660
DB	617 PVLVRVKAIMEFNPFEVARNVPECNDQVYVKGKAGEVRVCLHVQKSTRDLRREGQISVVT 676
QY	661 YDLALDSGRHSRAVFNFTKNSRTOVLGLTCTCETLKLQLPNCIEDPVPVILRLNF 720
DB	677 YDLALDSGRHSRAVFNFTKNSRTOVLGLTCTCETLKLQLPNCIEDPVPVILRLNF 736
QY	721 SLVGTPLSAFQNLRPVLAEDAQRIFLTALFPFEKNCNGNDNICQDDLSITPFSMLDCLVWG 780
DB	737 SLVGTPLSAFQNLRPVLAEDAQRIFLTALFPFEKNCNGNDNICQDDLSITPFSMLDCLVWG 796

QY	781 GPREFNVTTVTRNDGSDSYTQVTFPPDLVSRYKSTLQNSORSWRILACESASSTEV 840
DB	797 GPREFNVTTVTRNDGSDSYTQVTFPPDLVSRYKSTLQNSORSWRILACESASSTEV 856
QY	841 SGALKSTSCSINHIPIFENSEVFNFITFDVDSKASLGNKLLKANVTSENNMPRTNTEF 900
DB	857 SGALKSTSCSINHIPIFENSEVFNFITFDVDSKASLGNKLLKANVTSENNMPRTNTEF 916
QY	901 QLELPVXYAVVMVTVSHGVSTKYLNFPTASENTSRVMQHOYQVSNLQORSLSPLVFLVPV 960
DB	917 QLELPVXYAVVMVTVSHGVSTKYLNFPTASENTSRVMQHOYQVSNLQORSLSPLVFLVPV 976
QY	961 RLNCTVLDWRPQVTFSSNLSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCQRIQCDIP 1020
DB	977 RLNCTVLDWRPQVTFSSNLSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCQRIQCDIP 1036
QY	1021 PFGIOEBFNATLKNLSFDWYIKTSHNHLIVSTABEILFNDVSFETLLPGQCAFVRSQTEI 1080
DB	1037 PFGIOEBFNATLKNLSFDWYIKTSHNHLIVSTABEILFNDVSFETLLPGQCAFVRSQTEI 1096
QY	1081 KVBEFEPVNPPLLVGVSSVGGLLLELALITAAALYKLGFPKQYKDMMSGGPPGAEPPQ 1137
DB	1097 KVBEFEPVNPPLLVGVSSVGGLLLELALITAAALYKLGFPKQYKDMMSGGPPGAEPPQ 1153
RESULT 7	
ID	AA04136 standard; protein; 1153 AA.
XX	AA04136;
DT	25-MAR-2003 (revised)
DT	07-SEP-1990 (first entry)
XX	Alpha subunit of Mac-1 leukocyte adhesion receptor.
XX	Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
XX	non-specific defence system; integrin gene superfamily.
XX	Synthetic.
Key	Location/Qualifiers
FT	Region 1..16
FT	/label= signal peptide
FT	Modified-site 86..88
FT	/label= putative N-glycosylation site
FT	Modified-site 240..242
FT	/label= putative N-glycosylation site
FT	Modified-site 391..393
FT	/label= putative N-glycosylation site
FT	Modified-site 469..471
FT	/label= putative N-glycosylation site
FT	Modified-site 693..695
FT	/label= putative N-glycosylation site
FT	Modified-site 697..699
FT	/label= putative N-glycosylation site
FT	Modified-site 735..737
FT	/label= putative N-glycosylation site
FT	Modified-site 802..804
FT	/label= putative N-glycosylation site
FT	Modified-site 881..883
FT	/label= putative N-glycosylation site
FT	Modified-site 901..903
FT	/label= putative N-glycosylation site
FT	Modified-site 912..914
FT	/label= putative N-glycosylation site
FT	Modified-site 941..943
FT	/label= putative N-glycosylation site
FT	Modified-site 947..949
FT	/label= putative N-glycosylation site
FT	Modified-site 979..981
FT	/label= putative N-glycosylation site
FT	Modified-site 994..996

FT	Modified-site	/label= putative N-glycosylation site	421	QNTGWESNANVKQTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYTYEQTGGQSVCP	480
FT	Modified-site	/label= putative N-glycosylation site	437	QNTGWESNANVKQTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYTYEQTGGQSVCP	496
FT	Modified-site	/label= putative N-glycosylation site	481	PRGQARWQCDVLYGEGQPGWGFAGALTVDGVNGDKLTDVAIGAPBEDNKGAVYLF	540
FT	Modified-site	/label= putative N-glycosylation site	497	PRGQARWQCDVLYGEGQPGWGFAGALTVDGVNGDKLTDVAIGAPBEDNKGAVYLF	556
FT	Modified-site	/label= putative N-glycosylation site	541	HGTSGSGISPSHSORIKAGSKLSPRLQVFGQSLSGQDLTMDGLVLDLTVGAQGHVLLRSQ	600
FT	Region	/label= putative N-glycosylation site	557	HGTSGSGISPSHSORIKAGSKLSPRLQVFGQSLSGQDLTMDGLVLDLTVGAQGHVLLRSQ	616
XX	BP364690-A.	/label= putative transmembrane region	601	PVLRVKAIMFNPFEVARNFECDNDQVVKGEARVRLVQKSTRDLREGQIQSVVT	660
XX	25-APR-1990.		617	PVLRVKAIMFNPFEVARNFECDNDQVVKGEARVRLVQKSTRDLREGQIQSVVT	676
XX	17-AUG-1989;	89EP-00115159.	661	YDLALDSGRPHSRVAFNETKSTRQTVGLGTCTETKLQLPNCIEDPVSPIVLRNF	720
XX	23-AUG-1988;	88US-00235353.	677	YDLALDSGRPHSRVAFNETKSTRQTVGLGTCTETKLQLPNCIEDPVSPIVLRNF	736
XX	09-MAR-1989;	89US-00321239.	721	SLVGTFLSAFGNLRPVLAEDAQRLLFTALPPEKKNCGNDNICODDELSITFSPMSLDCLVVG	780
XX	(DAND) DANA FARBER CANCER INST INC.		737	SLVGTFLSAFGNLRPVLAEDAQRLLFTALPPEKKNCGNDNICODDELSITFSPMSLDCLVVG	796
XX	Springer TA, Corbi A;		781	GPREFNVTVTVNDGEDSYRTQVTFPPDLDSYKRVSTLONORSORSNRLACESASSTEV	840
XX	WPI; 1990-125938/17.		797	GPREFNVTVTVNDGEDSYRTQVTFPPDLDSYKRVSTLONORSORSNRLACESASSTEV	856
XX	N-PSDB; AAQ04043.		841	SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSNNMPTNKTEF	900
XX	New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating		857	SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSNNMPTNKTEF	916
XX	inflammation and viral infections, and in diagnosis.		901	QLELPVKYAVVMVTVSHGVSTKYLNFETASNTSRVMQHQYOVNSLQORSLPISLFLVFPV	960
XX	Disclosure; Page ?; -pp; English.		917	QLELPVKYAVVMVTVSHGVSTKYLNFETASNTSRVMQHQYOVNSLQORSLPISLFLVFPV	976
XX	Mac-1 alpha subunit is involved in the response to inflammation, i.e.		961	RLNQTVINDRPOVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVNVNCSTAVCQRIQC	1020
XX	recognition of and migration to sites of inflammation. It also attaches		977	RLNQTVINDRPOVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVNVNCSTAVCQRIQC	1036
XX	to cellular substrates as part of this function making it useful in		1021	FFGIOEEFNATLKGMLSPDWYIKTSHNHLIIVSTAEILFNDSDVFTLLPCOGAFVRSQTET	1080
XX	visualising endothelial tissue. Mac-1 is a member of the Integrin Gene		1037	FFGIOEEFNATLKGMLSPDWYIKTSHNHLIIVSTAEILFNDSDVFTLLPCOGAFVRSQTET	1096
XX	superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25		1081	KVEPFEVNPPLIIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEPO	1137
XX	-MAR-2003 to correct PA field.)		1097	KVEPFEVNPPLIIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEPO	1153
XX	Sequence 1153 AA;		RESULT 8		
XX	Query Match	99.78; Score 5858; DB 2; Length 1153;	AA07120		
XX	Best Local Similarity	99.66; Pred. No. 0;	ID	AA07120 standard; protein; 1163 AA.	
XX	Matches 1132; Conservative	3; Mismatches 2; Indels 0; Gaps 0;	XX	AA07120;	
QY	1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQIIVAAANQSGSLYQCDYSTGSCPEI	60	AC	25-MAR-2003 (revised)	
DB	17 FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQIIVAAANQSGSLYQCDYSTGSCPEI	76	DT	05-FEB-1991 (first entry)	
QY	61 RLQVPVEAVNMSLGLSLAATTSPPQLACGTVHQTCTSENTRYKGLCFGLGSLNRQOPQK	120	XX	p150.95 alpha subunit encoded by clone lambdaX47.	
DB	77 RLQVPVEAVNMSLGLSLAATTSPPQLACGTVHQTCTSENTRYKGLCFGLGSLNRQOPQK	136	XX	p150.95 leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia;	
QY	121 FPEALRGCPQSDSDIAFLIDGSGSIIIPDFRRMKFVSTVMEQLKSKTILFSLMOYSEEP	180	XX	rhinovirus.	
DB	137 FPEALRGCPQSDSDIAFLIDGSGSIIIPDFRRMKFVSTVMEQLKSKTILFSLMOYSEEP	196	XX	Synthetic.	
QY	181 RIHFTPFKFNPNPRLSKVITQLLGRTHATGVRKVIRELLNITNGARKNAPKILIVI	240	XX	Key	Location/Qualifiers
DB	197 RIHFTPFKFNPNPRLSKVITQLLGRTHATGVRKVIRELLNITNGARKNAPKILIVI	256	XX	Region	1..19
QY	241 TDGKFGDPLGYEDVPEADREGVIRVIGVDGAFRSEKSRQELNTIASKPPRDHVPQVN	300	XX	Region	/label= signal peptide
DB	257 TDGKFGDPLGYEDVPEADREGVIRVIGVDGAFRSEKSRQELNTIASKPPRDHVPQVN	316	XX	Region	20..44
QY	301 NFEALKTQNLREKI FAIEGTQSGSSSFHEMSQSGFSAAITNSGPLLSTVGSYDWAG	360	XX	Modified-site	61..63
DB	317 NFEALKTQNLREKI FAIEGTQSGSSSFHEMSQSGFSAAITNSGPLLSTVGSYDWAG	376	XX	/label= glycosylation site	
QY	361 GVFLYTSKSTPTNMTVDSDMDAYLGVAALILNRVQSLVGLGAPRYCHIGLVAMFR	420			
DB	377 GVFLYTSKSTPTNMTVDSDMDAYLGVAALILNRVQSLVGLGAPRYCHIGLVAMFR	436			

FT	Modified-site	89..91	/label= glycosylation site
FT	Modified-site	385..387	/label= glycosylation site
FT	Modified-site	392..394	/label= glycosylation site
FT	Modified-site	697..699	/label= glycosylation site
FT	Modified-site	735..737	/label= glycosylation site
FT	Modified-site	899..901	/label= glycosylation site
FT	Modified-site	904..906	/label= glycosylation site
FT	Modified-site	939..941	/label= glycosylation site
FT	Modified-site	1050..1052	/label= glycosylation site
FT	Domain	1108..1133	/label= transmembrane
XX			
PN	WO9010646-A.		
XX			
PD	20-SEP-1990.		
XX			
PF	09-MAR-1990;	90WO-US001257.	
XX			
PR	09-MAR-1990;	90WO-US001257.	
XX			
PA	(DAND) DANA FARBER CANCER INST INC.		
XX			
PI	Corbi AA, Springer TA;		
XX			
DR	WPI; 1990-304985/40.		
DR	N-PSDB; AAQ06068.		
XX			
PT	Treatment of viral esp. rhino-viral infection - by admin. of alpha sub-		
PT	unit of p150.95 cell surface adhesion receptor, opt. together with a beta		
PT	chain of CD-18 family.		
XX			
PS	Disclosure; Fig 3; 59pp; English.		
XX			
CC	Clone lambda X47 was isolated from a cDNA library constructed from total		
CC	RNA extracted from phorbol myristate acetate stimulated HL-60		
CC	myelomonocytic cells. The library was screened with oligonucleotide		
CC	probes based on tryptic peptide fragments of p150.95. The sequence can be		
CC	attached to appropriate control elements and expressed in prokaryotic and		
CC	eukaryotic cells. The protein can be used to treat or prevent rhinoviral		
CC	infection because it interacts with ICM-1 and inhibits cell-virus		
CC	attachment. It can also be used as an anti-inflammatory agent. See also		
CC	AAQ06063-4, AAQ06068, AAR07121-8 and AAR07152-6. (Updated on 25-MAR-2003		
CC	to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)		
XX			
SQ	Sequence 1163 AA;		
	Query Match	59.0%; Score 3464; DB 2; Length 1163;	
	Best Local Similarity	61.1%; Pred. No. 2.5e-279;	
	Matches 690; Conservative 138; Mismatches 295; Indels 6; Gaps 4;		
QY	1 FNLDENAFQBNARGFGQSVVQLQGRVWVGAPOEIVAANQGRSLYQDYSTGSCBPI 60		
DB	20 FNLDTELTAFRVDGSDVQVYANVWVGAPOKITAANQGTGLYQGYSTGACBPI 79		
QY	61 RLQVPVEAVNMSIGLSLAATTSPPOLLACGPTVHQCSENTYVKGICFLFGSNLRQPOK 120		
DB	80 GLQVPEAVNMSIGLSLAATTSPPOLLACGPTVHQCSENTYVKGICFLFGSNLRQPOK 137		
QY	121 PFEALRGCPQEDSDIAFLDGGSGIIPHDFRMKPEFVSTVMEQLKSKTLPFLMQYSEF 180		
DB	138 LPVSRQECFQEQEDIVFLDGGSGISSRNFAFMNFAVISOQFQRPSTQFSLMQFSNKF 197		
QY	181 RIHFTKGFQNNPNSLKPITOLLGRTHATGVKVRBELNITNGARKNAFKILIVI 240		

RESULT 9
AAW65091
ID AAW65091 standard; protein; 1163 AA.
XX
AC AAW65091;
XX
DT 28-SEP-1998 (first entry)
XX
DE Human Beta-integrin CD11c subunit protein.
XX

KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
 KW type-1 diabetes; atherosclerosis; multiple sclerosis; asthma;
 KW lung inflammation; acute respiratory distress syndrome; CD11c subunit;
 KW rheumatoid arthritis.
 XX Homo sapiens.
 XX US728533-A.
 XX 17-MAR-1998.
 XX 07-JUN-1995; 95US-00485618.
 XX 23-DEC-1993; 93US-00173497.
 XX 05-AUG-1994; 94US-00286889.
 XX 21-DEC-1994; 94US-00362652.
 XX (ICOS-) ICOS CORP.
 XX Van Der Vieren M, Gallatin WM;
 XX WPI; 1998-206565/18.
 XX Screening assay for modulators of integrin binding - using immobilised or
 XX labelled alpha-d polypeptide, useful for, e.g. treating type-1 diabetes.
 XX Example 5; Fig 1A-D; 106pp; English.
 XX This sequence represents a human beta-integrin CD11c subunit which is
 XX used to describe a method for identifying compounds that modulate the
 XX interaction of the beta-integrin alpha-d subunit with a binding partner
 XX of alpha-d which involves contacting an alpha-d polypeptide with an alpha
 XX -d binding partner, one of which is immobilised and the other of which is
 XX labelled, in the presence of a test compound, and determining if the
 XX compound affects binding between the alpha-d polypeptide and alpha-d
 XX binding partner, where the alpha-d polypeptide is alpha-d or its fragment
 XX comprising the cytoplasmic, transmembrane or extracellular domain of
 XX alpha-d. Compounds that modulate alpha-d binding could be used to treat
 XX diseases such as type-1 diabetes, atherosclerosis, multiple sclerosis,
 XX asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
 XX and rheumatoid arthritis
 XX SQ Sequence 1163 AA;
 Query Match 58.7%; Score 3450; DB 2; Length 1163;
 Best Local Similarity 61.0%; Pred. No. 3.8e-278;
 Matches 689; Conservative 144; Mismatches 290; Indels -6; Gaps 4;
 QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVYVGAPOEIVAAHQGLYQCDYSTGSCPEI 60
 DB 20 FNLDTEELAFRVDSAGFDSVVQYANSVWVGAPOKILAAQIGLYQGYSTGACPEI 79
 QY 61 RLOVPEAVNMSGLSLAATTSPPOLLACGPTVHQCSENYTKGLCLFLPGSNLRQPOPK 120
 DB 80 GLOVPEAVNMSGLSLAATTSPPOLLACGPTVHCEGRNMYLTGLCLGPT--QLTOR 137
 QY 121 FPEALRGCEDESDIAFLDGGSIIPHPFRMKFVTVMEQLKSKTFLSLMOSSEF 180
 DB 138 LPVSRCECPQEQDIFVLIDGSGISRRNPATMNFRAVISQFQFSLMQFSNKF 197
 QY 181 RIHFTKBPQNNPNSLVKPTQLGRTHATGVRKVRRELLNITNGARKNAFKILIVI 240
 DB 198 QTHFTFEERFRTSNPLSLASVHLQCFYVTAIQAIVVHRLFHASYGARRDAIKILIVI 257
 QY 241 TDCEKCDPLGYEDVPEADREGVIRVIGVGAFAFSEKSRQBLNTIASKPRDHVPQW 300
 DB 258 TDCKKEDSLDYNDVIMPDAAGIIRVAIGVGLAFQNRNSWKEINDIAKPSHELFKYE 317
 QY 301 NFELAKTIQNLREKIFAIBGTQTGSSSEFHEMSQEGFSAATISNGPLLTSTVGSDWAG 360
 DB 318 DFDALDQNLREKIFAIBGTETISSSEFELEMAQEGFSAVTPDPGLGAVGFTWSG 377
 QY 361 GVFLYTSKESKSTINMTRVDSMDNDAYLGVAAILLRNVQSLVGLAPRYQHIGLVAMFR 420

DB 378 GAFLYPNMSPTFINMSQENVDMRDSYLGYSSTELALWKGVQSLVGLAPRYQHIGKAVIFI 437
 QY 421 QNTGWESSNANVKCTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEOTRGQGVSVCP 480
 DB 438 QVSRQWRKAEVIGTIGSYFGASLCSVDVDSNGSTDLVLIGAPHYEOTRGQGVSVCP 497
 QY 481 PRGORARWQCDVAVLYGEGQGPWGRFGAALTIVLGVNVDGKLTVDVAIGAPGBEDNKGAVYLF 540
 DB 498 PRGWR-RWMCDAVLYGEGQGPWGRFGAALTIVLGVNVDGKLTVDVAIGAPGBEDNKGAVYLF 556
 QY 541 HGTSGSGISPSHSORIASGSLSPLOYFGQSLSGQDLTMDGLVDTLVGAQHVLLRSQ 600
 DB 557 HGVLPSPISPSHSORIASGSLSPLOYFGQSLSGQDLTMDGLVDTLVGAQHVLLRSQ 616
 QY 601 FVLRVGAIMEFNPREVARNVPCNDQVYVKGAEVVRVCLHVQKSTRDLREGQIQSVVT 660
 DB 617 FVLAVGVSMQFIPAEIIPSAFECREQVVSQTLVQSNICLYIDKRSKNLLSGRDLQSSVT 676
 QY 661 YDLALDSGRPHSRVAFNETKNSRTOVQLGLTOTCETKLQKLPNCLEDPVSPVLRNF 720
 DB 677 LDALAPGLSPRAIFQETKRSLSRVRLGLKAHCENFNLLPSCVEDSVIPILRLNF 736
 QY 721 SLVGTSLAPGNLRPVLAEADAQRLFTALFPPEKNCNDNICODDLISITFSFMSLDCLVVG 780
 DB 737 TLVGPFLAFNLRLPMLAALAQRYFTASLPKNCAGADHICQDNLGISFPGKLSLVG 796
 QY 781 GPRFNVTVVRNDGEDSYRTQVTFPPPLDLSYKRVSTLQNRQSRWRLACESASSTEV 840
 DB 797 SNLEINARVMVWNDGEDSYRTQVTFPPPLDLSYKRVSTLQNRQSRWRLACESASSTEV 854
 QY 841 SGALKSTSCSNHPIFPENSESVTNITFDVDSKASLGNKLLKXNTSENNMPTNKTEF 900
 DB 855 SQGTWSTSCRINHLIFRGGAQITFLATFVSPKAVGLDRLLLIANVSSENNIPRTSKIF 914
 QY 901 QLELPKYAVYVMTVSHGVSTKYLNFTAS-ENTSRVMOHQVQVNLGORSIPISLVFLVP 959
 DB 915 QLELPKYAVYVMTVSHGVSTKYLNFTAS-ENTSRVMOHQVQVNLGORSIPISLVFLVP 974
 QY 960 VFLNCTVMDRPOVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCISAVCORIQCDI 1019
 DB 975 VELNQEAVMMDVEVSHPNPDLRCSSEKIAPASDPLAHIQKNFVLDLSIACCLAFRCDV 1034
 QY 1020 PFFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSTVFTLLPQOGAFVRSQTE 1079
 DB 1035 PSFSVQEBLDTLKNLSFQWVRQILQKKVSVSWAEIIFDTSVYSQLPQOEAFAFRAQTI 1094
 QY 1080 TKVEPFVFNPLPLIVGSSVGLLALITATYALYKLGFEKQYKXMMSE 1128
 DB 1095 TVLEKYKVNPIPLIVGSSIGGLLLALITATYALYKLGFEKQYKXMMSE 1143
 RESULT 10
 AAB07361
 ID AAB07361 standard; protein; 1163 AA.
 XX
 AC AAB07361;
 XX
 DT 17-JAN-2001 (first entry)
 XX
 DE Human CD11c protein sequence.
 XX
 KW Human; macrophage infiltration inhibition; alpha_d integrin;
 KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;
 KW inflammation; leukocyte adhesion deficiency; IAD; Type 1 diabetes;
 KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
 KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
 KW rheumatoid arthritis; central nervous system injury; CD11c.
 XX Homo sapiens.
 OS
 XX W0200029446-A1.
 PN
 XX

PD 25-MAY-2000.
XX
PF 16-NOV-1999; 99WO-US027139.
XX
PR 16-NOV-1998; 98US-00193043.
PR 08-JUL-1999; 99US-00350259.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gallatin MW, Van Der Vieren M;
XX
DR WPI, 2000-387751/33.
XX
XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT macrophage infiltration and reduce inflammation at central nervous system
PT injury sites.
PS Example 5; Fig 1; 270pp; English.
XX
CC Integrins are a class of membrane-associated molecules that participate
CC in cellular adhesion. Integrins are made up of an alpha subunit and a
CC beta subunit. One class of human integrins are restricted to expression
CC in white blood cells and have a common beta2 subunit: the leukocyte
CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
CC have an important role in immune and inflammatory responses. The present
CC protein sequence is the human integrin alpha subunit CD11c. This sequence
CC was used in an alignment to identify a novel beta2 integrin alpha
CC subunit: alpha d (AA60014 and AA607359). The present sequence has
CC approximately 66% identity to the protein sequence of alpha.d. The
CC Alpha d gene and protein may be useful in therapy for diseases linked to
CC aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple
CC sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
CC distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
CC (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the
CC inhibition of macrophage infiltration at the site of a central nervous
CC system injury. The monoclonal antibodies can also be used to detect and
CC diagnose Crohn's disease
XX
SQ Sequence 1163 AA;
Query Match 58.7%; Score 3450; DB 3; Length 1163;
Best Local Similarity 61.0%; Fred. No. 3.8e-278;
Matches 689; Conservative 144; Mismatches 290; Indels 6; Gaps 4;
QY 1 FNLDTENAMTQENARGGQSVVQLQSGRVVVGAPQEIIVAAHQESLYCCDYSGSCBPI 60
DB 20 FNLDTEELTAPEVDSAGGDSVVQYANSWVVGAPQKLIIRANQIGELYCCGYSTGACBPI 79
QY 61 RLOVPEAVNMSLGLSLAATTPPOLLAGCTVHTCTSENVTYKGLCFPLRSNLRQDPK 120
DB 80 GLOVPEAVNMSLGLSLAATTPPOLLAGCTVHTCTSENVTYKGLCFPLRSNLRQDPK 137
QY 121 PPEALRGCPQEDSDAFLIDGSGSIIPHDFFRMKEFVSTVMEQLKSKTFLSLMOYSREF 180
DB 138 LPVSQCEPQEQDIVFLIDGSGSISSRNFAFMVAVISQFQPSQFSLMOYSNKP 197
QY 181 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGKVRKVIKRELLNITNGARKNAFKILIVI 240
DB 198 QTHFTFEBFRSTNPLSLASVHQLQGGFTYATATQNVVHRLFRASYGARRDAIKILIVI 257
QY 241 TGGKPGPLGVEDVPEADRGVRYVGVGDAPFSEKSRQELNTIAKPEPRDHVPOVN 300
DB 258 TGGKPGPLGVEDVPEADRGVRYVGVGDAPFSEKSRQELNTIAKPEPRDHVPOVN 317
QY 301 NFEALKTIQNLREKIFAEGTQSSSSFEHMSQEGFSAATSGNPLLTGVSQYDAG 360
DB 318 DFDALDKIQNLKEKIFAEGTQSSSSFEHMSQEGFSAATSGNPLLTGVSQYDAG 377
QY 361 GVFLTYSKEKSTFNNTRVDSMDNAYLYGAAAIILNRVQSIATLGAPYOHIGLVAMFR 420
DB 378 GAFLYPPNMSPTFINNSQENVDMDRSYLGSTELALWKGVSQSLVGLAPYOHIGKAVIFI 437
QY 421 QNTGHWESNANVKGQIIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTREGGQSVCP 480

DB 438 QVSRQMRMKAETVGTQIGSYFGASLCSVDVDTDGSSTDLVLIGAPHYEQTREGGQSVCP 497
QY 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDAVAGAPGEDNNGAYILF 540
DB 498 PRGWR-RWNCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDAVAGAPGEDNNGAYILF 556
QY 541 HGTSGSGISPSHSORTAGSKLSPRLQYFGOSLGGDLTMDGLVTLVGAQGHVLLRSQ 600
DB 557 HGVLPSPISPSHSORTAGSKLSPRLQYFGOSLGGDLTMDGLVTLVGAQGHVLLRSQ 616
QY 601 PVLRVKAIEMFNEPREVARNVFECDQVWKGKAGEVRVCLHVQKSTFDRLEBEOIQSVT 660
DB 617 PVLVWVSMQFIPAEIPRSFAFECEQVWSEQLTVQSNICLYIDKRSKLLSGSRDLQSVT 676
QY 661 YDLALDSGRPHSAVFNENKSTRTQVGLTQCTETFLKQLPNCIEDVSPVTVRLNF 720
DB 677 LDALAPGRSLSPRAIQETKRSLSRVRLGLKAHCENFNLLPSCVEDSVIPIILRLNF 736
QY 721 SLVGTPLSAFGRNLRPVLAEADAQRLFTALPFPEKXNCGNDNICQDDLSITPFMSDLCI 780
DB 737 TLVGKPLAFNLRPMLAALAQRYFTASLPFEKXNCGADHICQDNLGISFSPGLKSLVG 796
QY 781 GPREFNVTVVRNDGDSYQVTPPPDLSTRKUSTLQNSORSWRSLACESASTEV 840
DB 797 SNLELNAEVMVMNDGDSYCTTITFSHPAGLSYRYVAEGQKQGLRSLHLTC--CSAPVG 854
QY 841 SGALKSTCSINHIPIPPENSEVTEFITFDVDSKASLGNKLLKANVTSENMPRTNKTET 900
DB 855 SGTWSTSCINHLIFRGQAIQIFLTFDPSKAVGLDRLLANVTSENMPRTNKTET 914
QY 901 QLELPVYAVVYVWVTSKGVSKYLNFTAS-ENTSRVMOHQYQVSNLQORSLPISLVLPV 959
DB 915 QLELPVYAVVYVWVTSKGVSKYLNFTAS-ENTSRVMOHQYQVSNLQORSLPISLVLPV 974
QY 960 VRLNQTIVDRPQVTFSENLSSCTCHTKERLPSSDFLAELKAPVWNCSTAVCORICDI 1019
DB 975 VELNQEAVMVMDEVSHPFNSLRCSEKAPPSADDFLAHQKQVLPDSCSACGRFCDV 1034
QY 1020 PFGIQEENATLKGNSLFDWYINTSHNLLIYVSTABIFENDSVFTLLPQGAFFRSQTE 1079
DB 1035 PSFSVQSELDFTLKGNSLFGWVRQILQKKYVSVVABIIEDTSVYSQLPQGAFFRAQTI 1094
QY 1080 TKVPEPVPNPLTVGSSVGGELLALLIITAAALYKLGFFKRYKDMSE 1128
DB 1095 TVLEKYKHNPPIPIVSGSSIGGLLLALITAVLYKVGFFKRYKEMMEB 1143
RESULT 11
ABG61470
ID ABG61470 standard; protein; 1163 AA.
XX
AC ABG61470;
XX
XX 27-AUG-2002 (first entry)
XX Human Beta2 integrin alphaCD11c subunit.
XX
XX Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit; LAD;
XX leukocyte adhesion deficiency; inflammatory response; diabetes;
XX multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
XX inflammatory bowel disease; Crohn's disease; ulcerative colitis;
XX immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
XX intracellular cell adhesion molecule; vascular cell adhesion molecule;
XX locomotor recovery; locomotor damage; locomotor impairment;
XX autonomic dysfunction; sensory dysfunction; spinal cord injury.
OS Homo sapiens.
XX
XX WO200230980-A2.
XX
XX 18-APR-2002.
XX

PF	15-OCT-2001; 2001WO-US032059.	Db	438	QVSRQWRKAEVIGTQIGSYFGASICSVDVDTGSDTDLVLIGAPHYVQTRGGQVSVCP	497
XX					
PR	13-OCT-2000; 2000US-00688307.	Qy	481	PRGQARWQCDVLYGEGQCPWGFAGALTVDVNGDKLTDVVAIGAGEENRGA	540
XX					
PA	(ICOS-) ICOS CORP.	Db	498	PRGMR-RWCDVLYGEGQCPWGFAGALTVDVNGDKLTDVVAIGAGEENRGA	556
XX					
PI	Gallatin WM, Van Der Vieren M;	Qy	541	RGTSGLSPSHSQRISAGSLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQHVL	600
XX					
XX	WPI; 2002-463260/49.	Db	557	HGVLGPSISPSHSQRISAGSLSPRLQYFGQSLSGQDLTMDGLVDLTVGA	616
DR					
XX					
XX	Use of an anti-alpha-d monoclonal antibodies for promoting locomotor	Qy	601	PVLRVKAWEFNPREVARNFECNDQVVGKGEASRVRVCLHVQKSTRDLREGQ	660
PT	recovery, inhibiting locomotor damage, limiting locomotor impairment, or				
PT	limiting autonomic and sensory dysfunction following spinal cord injury.	Db	617	PVLWVGVMQFIPAEIPRSAFECEQVYSEQLVQSNICLYIDKESKLLSRDQSS	676
XX					
PS	Example 5; Page 194-198; 270pp; English.	Qy	661	VDLALDSGRPHSRVFNETKSTRTOVLTGTCTETKQLQCPNCIEDPVPVLRNF	720
XX					
CC	The invention relates to promoting locomotor recovery, inhibiting	Db	677	LDLALAPGRLSPRALFOETKNSLSRVAVLGKAKHCENFLLPSCVEDSVIPI	736
CC	locomotor damage, limiting locomotor impairment, or limiting autonomic				
CC	and sensory dysfunction following spinal cord injury by administering an	Qy	721	SLVGTPLSAGNLRPVLABDAORLEPTALPPFKKNGNDNICODDLISITFSP	780
CC	anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a				
CC	spinal cord injury victim. The method also involves the use of a ligand	Db	737	TLVKGKPLAFRLRPLAALAQRYFTASLPFKKNGADHICODNLGISFSPGLK	796
CC	selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule,				
CC	vascular cell adhesion molecule). The method is useful for promoting	Qy	781	GPREEFNVTYVRNDGEDSVRTQVTFPPFLDLSYKRVSTLQNRQSRWLACES	840
CC	locomotor recovery, inhibiting locomotor damage, limiting locomotor				
CC	impairment, or limiting autonomic and sensory dysfunction following	Db	797	SMLENAEVMWWDGEDSYGTITFESHAGLSYRVASGQKQGLRSLHLC--CS	854
CC	spinal cord injury. In particular, the spinal cord injury comprises				
CC	compression of the spinal cord. The antibodies are also useful for	Qy	841	SGALKSTSCSINHPIPPENSEVTFNITFDVSKASLGNKLLKXANTSENMPRT	900
CC	reducing inflammation at the site of a central nervous system injury. The				
CC	specification also details the identification of Beta2 integrin alpha	Db	855	SQGTWSTSCRINHLIFRGAQITFLATFDVSPKAVGLDELILLIANVSSENN	914
CC	cDNAs and proteins, for use in raising the antibodies. Beta2 integrins				
CC	are implicated in diseases such as IAD (leukocyte adhesion deficiency,	Qy	901	QLELPVKYAVYVAVTSHGVSTKYLNTAS-ENTSRVMOHYOVSNLQORSLP	959
CC	inflammatory response, diabetes, multiple sclerosis, arthritis, graft				
CC	atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative	Db	915	QLELPVKYAVYVIVSSHQFTKYLNFSSSEKESHVAVHRYVNNLQORDLP	974
CC	colitis, immune complex alveolitis and leukaemia. The present sequence is				
CC	a Beta2 integrin alpha subunit sequence included for comparison with the	Qy	960	VELNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNVNCIA	1019
CC	Beta2 integrin alpha protein sequences				
XX		Db	975	VELNQEAVMWVDESHPPQNPSPSLRCSSEKIAPASDPLAHIQKNPVLD	1034
SQ	Sequence 1163 AA;				
	Query Match 58.7%; Score 3450; DB 5; Length 1163;	Qy	1020	PFPGTQEBFNATLKGNSLDFWIKTSHNHLIVSTAEILFNDVSFTLLP	1079
	Best Local Similarity 61.0%; Pred. No. 3.8e-278;				
	Matches 689; Conservative 144; Mismatches 290; Indels 6; Gaps 4;	Db	1035	PSFSVQEELEDTFLKGNLSFGWVRQILQKVVSVVVAEILFDTSVYS	1094
Qy	1 FNLDTEAMFQENARFGQSVVLQGSRVVVGAPQIVANQEGSLVQDYTG	Qy	1080	TKVVEPEVNPPLIVGSSVGGILLALITAAALYKLGPFKQYKMMSE	1128
Db	20 FNLDTELTAFRVDGAGSDGWSQYANSVVVVGAPQKIIAANOIGGLY	Db	1095	TVLEKXKVENPIPLIVGSSIGGLLALITAVLYKVGPFKQYKMMSE	1143
Qy	61 RLQVPVAVNMSLGLSLAATTPPQLACGPTVHTCSENTYVKELCF	RESULT 12			
Db	80 GLQVPEAVNMSLGLSLASTTSPQLACGPTVHCEGRNMYLGLCFL	ABU07406			
Qy	121 FPEALRGCPQSDIAFLIDGSGSIIPHDFRMKEPSTVWEQLKSK	ID	ABU07406 standard; protein; 1163 AA.		
Db	138 LFSRQECPEQEDIVFLIDGSGSISSENFATMNFRAVISQF	AC	ABU07406;		
Qy	181 RIHTFEFONNPFSLVAPKITOLLGRTHATGVKRVIRELINTNG	XX			
Db	198 QTHTFEERTSNPLSLASVHGQGFYTTATAIQNVHFLPHASYG	DT	28-JAN-2003 (first entry)		
Qy	241 TDGKPGDGLGYEDVPEADREGVIRVYVGDAFRSEKSRQELNT	XX	Protein differentially regulated in prostate cancer #9.		
Db	258 TDGKKGSDLDYKDVIPMADAAGIIRYVGLAFQNRNSWKEIND	XX	Prostate cancer; gene expression; differential regulation;		
Qy	301 NFEALKTONLRREKIFAIBGTQTSQSSSEHEMSQSFSAITNS	XX	molecular marker; drug target; cancer detection; cancer diagnosis;		
Db	318 DFDALDKDQLQNKREKIFAIBGTETIISSSSELENAQEGFSA	XX	cancer staging; cancer grading; cancer assessing; cancer monitoring.		
Qy	361 GVFLYTSKSKSTFTNMTVRDSDMDNDVIGYAAAIILNRVQSLV	OS	Homo sapiens.		
Db	378 GAFYPPNMSPTFNMSQENVDMSYLGSTELALWKGVSILVGLAP	XX	WO200281638-A2.		
Qy	421 QNTGWESNANVKTQICAYFGASICSVDVDSNSTDLVIGAPHYV	PN	17-OCT-2002.		
		PD	08-APR-2002; 2002WO-US010824.		
		PF	06-APR-2001; 2001US-0281731P.		
		PR	06-APR-2001; 2001US-0281732P.		
		XX	(ORIG-) ORIGENE TECHNOLOGIES INC.		

XX Sun Z, Jay G;
XX MPI; 2003-058520/05.
XX Novel genes which are differentially regulated in prostate cancer, useful
XX for diagnosing prostate cancer in prostate tissue sample and assessing
XX therapeutic or preventive intervention in prostate cancer patients.
XX
XX Claim 1; Page 225-228; 416pp; English.
XX
XX The invention describes genes (I) which are differentially regulated in
XX prostate cancer. (I) is useful for diagnosing a prostate cancer in a
XX sample comprising prostate tissue, which involves determining the number
XX of target genes which are differentially-regulated in the sample, where
XX the number is indicative of the probability that the sample comprises
XX prostate cancer. (I) is useful for assessing a therapeutic or preventive
XX intervention in a subject having a prostate cancer, which involves
XX determining the expression levels in a sample comprising prostate tissue
XX of target genes which are differentially-regulated in prostate cancer.
XX preferably, the expression levels of at least 10 genes are determined.
XX (I) is also useful for identifying agents that modulate a biological
XX activity of a polypeptide differentially-regulated in prostate cancer
XX cells, which involves contacting a polypeptide differentially-regulated
XX in prostate cancer cells with a test agent under conditions effective for
XX the test agent to modulate a biological activity of the polypeptide, and
XX determining whether the test agent modulates the biological activity. (I)
XX is useful as molecular markers, as drug targets, and for detecting,
XX diagnosing, staging, grading, assessing, monitoring, prognosticating,
XX preventing or treating, determining predisposition to diseases and
XX conditions especially relating to prostate cancer. (I) and its expression
XX products are used in the diagnostic test to assay for presence of cancer
XX e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
XX blood etc. (I) is useful for assessing cancer e.g., to determine the type
XX of cancer, its stage of development, the nature of genetic defect, etc.
XX The polypeptide encoded by (I) can be used as target for therapy or drug
XX discovery. (I) can also be used for expressing the polypeptide and thus
XX for searching specific binding partners of the polypeptide. (I) is useful
XX in therapeutic applications to treat prostate cancer. The identification
XX of specific genes, and groups of genes, expressed in pathways
XX physiologically relevant to prostate cancer permits the definition of
XX functional and disease pathways and the delineation of targets in these
XX pathways which are useful in diagnostic, therapeutic, and clinical
XX applications. This is the amino acid sequence of a protein differentially
XX regulated in prostate cancer
XX
XX Sequence 1163 AA;
XX
XX Query Match 58.7%; Score 3448; DB 6; Length 1163;
XX Best Local Similarity 60.9%; Pred. No. 5.5e-278;
XX Matches 688; Conservative 136; Mismatches 299; Indels 6; Gaps 4;
XX
XX 1 FNLDTENAMTFOENARFGQSVVQGGSRVVVVGAPQEIIVAAHQSLYQCDYSTGSCPEI 60
XX 20 FNLDTEELTAFRVDSAGFDSVVQVANSVVVVGAPQKITAANQTCGLYQCGYSTGACEPI 79
XX
XX 61 RLQVPVEAVNMSLGLSLAATTPSPOLLACGPTVHOTCSENVYKGLCFGLFNSLRQPOK 120
XX 80 GLQVPEAVNMSLGLSLAATTPSPOLLACGPTVHOTCSENVYKGLCFGLGPT--QLTOR 137
XX
XX 121 FPEALRGCPQSDIAFLIDGSGSIIPDFRWMKBFVSTWMEQLKSKTFLFSLMOYSBEF 180
XX 138 LPVSRQECPRQGDIVFLIDGSGSISSRNFAFMNFRVAVISQFORPSTQFSLMQFSNKF 197
XX
XX 181 RIHFTKPEQNPNRSLVKPTTOLLGRTHWTATGVRKVTREILNITNGARKNAFKILIVI 240
XX 198 QTHLTPEEFRTSNPLSLASVHQLQGTYYTATIGNVVRHLPHASYGARRDATTILIVI 257
XX
XX 241 TDCGKPGDPLGVEDVTPRADRCVIRYVIGVDAPFSEKSRQELNTIASKPRDHVQVFN 300
XX 258 TDGKKGEDTLDYKDVIFMDADAGIRYALGVGLAQNRNSWELNDIASKPQEHFKVE 317
XX
XX 301 NFEALNTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGVSNMAG 360

Db 318 DFDALDKDIQTOLAREKIFPIETCTETSSSELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377
Qy 361 GVFLYTSKSKSTPINMTRVDSMDNDAYLGYAAAILLENRVQSLVGLAPRPHOHLVAMPR 420
Db 378 GAFLYPPNNSPTFINMSQENVDNRDSYLGYSTELALWKGVSQSLVGLAPRYOHTGKAVFT 437
Qy 421 QNTGMESNANVGTGTOIGAVFGASLCSVDVDSNGSTDLVLIGAPFYFYEQTRGGQSVCP 480
Db 438 QVSRQWRKRAEVTGTQIGSFSGSLCSVDVDSNGSTDLVLIGPFYFYEQTRGQSVCP 497
Qy 481 PRQARFQWCDAYLYGEGQCPWCRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYLF 540
Db 498 PRQWR-RWNCDAVLYGEGQHPWCRFGAALTVLGDVNGDKLTDVVIAGPGEENRGAYLF 556
Qy 541 HGTSGSGISPSHQRAGSLKSPLOYFGOSLGGGDLTWDGLVDTLVGQGHVLLRSQ 600
Db 557 HGVLFPSISPSHQRAGSLKSPLOYFGQALSGGDLTQDGLVGLAVGARGQVLLRTR 616
Qy 601 PVLEVKALMEFNEPREVARNVPECNDQVVGKEAGEVRVCLVHVKSTDRDRBQIOISVVT 660
Db 617 PVLWVGVMQFIPAEIPRSAFECEBQVVSQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676
Qy 661 YDLALDSGRPHSAVFNENSTRTOTVGLTQCTETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 LDIALDPERLSPRATFOETKRSLSRVRLUGLKAHCENFNLLPSCVEDSVPTILRNF 736
Qy 721 SLVGTPLSAGFNLRPVLAEQAQLFTALFFEXKNCNDNI CODDLSITTFPMSLDCLVWG 780
Db 737 TLVCKPELLAPNLRPMLAADAQRYFTASLFFEXKNCADHICQDNLGISFSFPGKSLVWG 796
Qy 781 GPRFNVTVVRNDGDSYTOTFPFPLDLSYRKVSTLQNRQSQRSLACASSTEV 840
Db 797 SNLENAEVMVMNDGDSYGTTFITFHPAGLSYRYVAEGQQLRSLHLTCDSPAVG-- 854
Qy 841 SGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPTNKTFF 900
Db 855 SQGTWSTSCINHLIFRGAQIFLATFDVSPRAVLGDRLLLTANVSENNTPRTSKTF 914
Qy 901 QLELPKYAVVMVTSYGSTKYNLTAS-ENTSRVMHQYOVQVSNLQORSIPISLVLVP 959
Db 915 QLELPKYAVVMVTSYSSHEQFTKYNLFSESEKESHVAMHRYQVNNLQORDLPVSNFVWP 974
Qy 960 VRLNQTVMDRPOVTFENSLSTCTKRLPSHSDFLAELRKAPVNCSTAVCQRIQDI 1019
Db 975 VELNQEAVMMDVEVSLPQNSLRCSSEKIAGPASDFLAHQKNPVLDCSIAGCLFRCDV 1034
Qy 1020 PFGIOEFENATLXGNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQCAFVRSQTE 1079
Db 1035 PFSVQBELDFTLKGNLSPGVVRQILQKYSVWSVASITEDTSVYSQLPQGEAFWRAQT 1094
Qy 1080 TKVPEFVFPNPLPIVGVSSVGGILLALLITAAKYKLGFPKQYKDMXSE 1128
Db 1095 TVLEKYVHNPTPLVIGSSIGGLLLALLITAVLYKVGFFKRYKEMMBE 1143
XX
XX RESULT 13
XX AAR78166
XX ID AAR78166 standard; protein; 1161 AA.
XX AC AAR78166;
XX DT 28-DEC-1995 (first entry)
XX DE Human beta-2 integrin alpha-d.
XX XX
XX KW Beta-2 integrin alpha-d subunit; antiinflammatory; arteriosclerosis;
XX inflammatory bowel disease; asthma.
XX OS Homo sapiens.
XX XX
XX FT Key Location/Qualifiers
XX Domain 17..1108

FT Region /note= "extracellular domain"
 FT 150..352
 FT /note= "this region is homologous to the insert common to
 FT Chila,b,c and may be a site for interaction with ICAM
 FT family proteins"
 FT 465..474
 FT Binding-site /note= "putative cation binding site"
 FT 518..527
 FT Binding-site /note= "putative cation binding site"
 FT 592..600
 FT Binding-site /note= "putative cation binding site"
 FT 1109..1128
 FT Region /note= "transmembrane region"
 FT 1129..1161
 FT Domain /note= "cytoplasmic domain"
 PN W09517412-A1.
 XX
 PD 29-JUN-1995.
 XX
 PF 21-DEC-1994; 94WO-US014832.
 XX
 PR 23-DEC-1993; 93US-00173497.
 PR 05-AUG-1994; 94US-00286889.
 XX
 PA (ICOS-) ICOS CORP.
 XX
 PI Gallatin WM, Van Der Vieren M;
 XX
 DR WPI; 1995-240603/31.
 DR N-PSDB; AAQ91712.
 XX
 PT Alpha sub-unit polypeptide of human beta 2 integrin - used to identify
 PT potential antiinflammatory agents, for the treatment of graft
 PT arteriosclerosis, inflammatory bowel disease, asthma, etc.
 XX
 PS Claim 7; Page 82-87; 172pp; English.
 XX
 CC -TMI was based on a partial cDNA clone (given in AAQ91727) of canine alpha
 CC -TMI was used to screen a human spleen cDNA library to identify clone
 CC 19A2 encoding the beta-1 integrin alpha-subunit, alpha-d. The cDNA was
 CC manipulated to allow expression of recombinant alpha-d subunit in COS and
 CC CHO cells
 XX
 SQ Sequence 1161 AA;
 Query Match 58.2%; Score 3417; DB 2; Length 1161;
 Best Local Similarity 59.8%; Pred. No. 2.1e-275;
 Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;
 QY 1 FNLDTENAMTFQENARGFQGSVVLQGSRRVVVVGAPQEIIVAAVNAQGSGLYQCDYSTGSCBPI 60
 DB 17 FNLDVEEPTIFQEDAGGFGQSVVQFGGSLVVGAPLEVVAAVNAQGRLYDCAANTGMCQPI 76
 QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTVYKGLCELPFGSLNQBPQPK 120
 DB 77 PLHIRPEAVNMSLGLTAAATNGSLRLLACGPTLHVRVCGENSYSKSGCLLGLSRW-EITQT 135
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPDFRMBKBFVSTVMEQLKSKTFLFSIMQYSEEF 180
 DB 136 VPDATPECHQEMDIVFLIDGSGSIDQDNDFNMKGQVQVMQVQFSGTDTLPALMQYSNLL 195
 QY 181 RHFTFKPQNPNPRLSKVPITOLLGRTHRTATGKRVIRELLNITNGARKNAKILLIVI 240
 DB 196 KIHFTFTQRTSPSQSLVDPIVLQKGLTFTATGTLTVVTFQPHKNGARKSAKILLIVI 255
 QY 241 TQGEKFGDPLGYEDVPIPEADRGVTRYVIGVGDADFSEKSRQELNTIASKPPDRHVFQVN 300
 DB 256 TDQQKYDPLEYSDVPIPEAKAGIIEYALGVGHAFQGPARTARQELNTISSAPPQDHVFKVD 315
 QY 301 NFEALATIONQUREKIPALTEGTQSSSFEHMSQEGFSAITNGPLLSVTGVDWAG 360
 DB 316 NEAALGSIQOLQEKIYAVEGTQSRASSSFQEMSQEGFSTALITMDGLFLGAVGFSWSG 375

QY 361 GVFLYTSKESKTFINMTRVDSMDNDAYGAAAIILNRRVQSLVGLGAPRYQHIGLVAMFR 420
 DB 376 GAFLYPNMSPPTFINMSQENVDMEDSYLGYSTELAMKGVQNLVLGAPRYQHTKAVIFT 435
 QY 421 QNTGMWESNANVKTGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTGSGGVSCPL 480
 DB 436 QVSRQWEKKAEVTGTQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYEQTGSGGVSCPL 495
 QY 481 PRQORAEWQCDVLYGQGGQPGWGFGAALTIVLGVNMGDKLTDVAIGAPGEDNFGAYLYF 540
 DB 496 PRGQVQWQCDVLYRGQGGHGWGFGAALTIVLGVNMGDKLTDVAIGAPGEDNFGAYLYF 555
 QY 541 HGTSGSISPSHSQRIAGSKLSRLQYFGSLSGGQDLTMDGLVDLTVGQGHVLLRSQ 600
 DB 556 HGASESGISPSHSQRIASSQLSRLQYFGQALSGGQDLTQDGLMDLAVGARGQVLLRS 615
 QY 601 PVLVVKAIMENPREVARNVFECDNDQVVKGEAGEVVRVCLHVQKSTRDRLEBQIQSVVT 660
 DB 616 PVLKVGVMRFPSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673
 QY 661 YDLALDSGRPHSRVAVNETKNSRRQTVGLTQCTETKLKLPNCIEDPVSPIVLINP 720
 DB 674 FDLALDPGLTSLRAIPNETKNPTLTRKTLGLGHCETLKLPLDPCVEDVVSPIHLHNF 733
 QY 721 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFPFKKNCNDNICODDLSTIF8FMSLDCLVVG 780
 DB 734 SLVREPIPSQNLRLPVLAVGQDLFTASLPFKKNCQDGLCEGDLGYTLSPSLQTLTVG 793
 QY 781 GPREFNTVTVRNDGDSYRTQVTFPPPLDLVYRKVSTLQNRQSRWRLACESASSTEV 840
 DB 794 SSLESLAVITVWNAGDSYGTWVSLYYPAGLSHRVSGAKQPHQSALRLACETV-PTED 852
 QY 841 SGALKSTSCSINHPIPPENSEVENTNITFDVDSKASLGNKLLKLLKANVTSENNMPTNTEF 900
 DB 853 EG-LRSSRCSVNHPIFHEGNGHFIVFDVSKATLDGRLMLRASSENKAKSSSKATP 911
 QY 901 QLELPVYAVYVTVSHGVSTKYNLP-TASENTSRVMQHVQVQVSNLQORSIPISLVFLVP 959
 DB 912 QLELPVYAVYVTVSHGVSTKYNLP-TASENTSRVMQHVQVQVSNLQORSIPISLVFLVP 971
 QY 960 VRLNQTVIMDRPQVTFSENLSSTCTKERLPSHSDFLAELRKAPVNCSTAVCORICDI 1019
 DB 972 VLLNGVAVVMDVMEAPQSLL--PCVSEKPPQHSDFLTQISRSPLMDCSADCLQFRCDV 1029
 QY 1020 PFGIQEEFNATLKNLSFDWYIKTSHNHLIYVSTAEILFNDVSFTLLPGQGAFFVRSQTE 1079
 DB 1030 PPSVQEEELDTLKNLSFGWVRETLQKVLVWVSAITEDTSVYSQLPQGAFFWRAQME 1089
 QY 1080 TKVPEFPVNPPLPLIVGSSVGLILLALITAAALYKLGFEKQYKDMKSE 1128
 DB 1090 MVLBEDEVYNAIPIIMGSSVGLILLALITATLYKLGFEKQYKDMKSE 1138
 RESULT 14
 AAW23049
 ID AAW23049 standard; protein; 1161 AA.
 XX
 AC AAW23049;
 XX
 DT 24-FEB-1998 (first entry)
 XX
 DR Human beta 2 integrin alpha d subunit.
 XX
 KW Beta 2 integrin alpha d subunit; human; cell migration; cell adhesion;
 KW phagocytosis; diabetes; atherosclerosis; multiple sclerosis; asthma;
 KW psoriasis; lung inflammation; acute respiratory distress syndrome;
 KW rheumatoid arthritis.
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FT Domain 17..1108

```

Domain
/label= Extracellular_domain
150..352
/Note= "region homologous to the I (insertion) domain
common to CD11a, CD11b and CD11c"
1109..1128
/label= Transmembrane_domain
/Note= "homologous to the human CD11c transmembrane
region"
1129..1161
/label= Cytoplasmic_domain

WO9731099-N1.
28-AUG-1997.
24-FEB-1997; 97WO-US002713.
22-FEB-1996; 96US-06605672.
XX (ICOS-) ICOS CORP.
XX Gallatin WM, Van Der Vieren M;
PI WPI; 1997-435154/40.
DR N-PSDB; AAT79220.
XX Hybridoma 199M and antibody secreted by it - specific for new rat beta2
PI integrin subunit, useful to detect subunit in cells and modulate its
PT activity.
XX
XX Example 5; Page 116-120; 222pp; English.
XX
XX This polypeptide comprises a novel human beta 2 integrin subunit,
CC designated alpha d. Its sequence was deduced from a cDNA clone (see
CC AAT79220) isolated from a spleen cDNA library. Alpha d is involved in
CC cell migration, phagocytosis and cell-cell interaction. Recombinant alpha
CC d polypeptides can be expressed in transfected host cells for use in
CC assays for identifying antibodies or other compounds that modulate alpha
CC d activity or which modulate the interaction between alpha d and a
CC ligand, for treating or preventing diseases in which macrophages are
CC implicated. Treatment is applicable to disease states in which alpha d
CC binding, or localised accumulation of cells which express alpha d, is
CC implicated such as, for example, type I diabetes, atherosclerosis, multiple
CC sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
CC distress syndrome and rheumatoid arthritis
XX
XX Sequence 1161 AA;
Query Match
Best Local Similarity 59.8%; Pred. No. 2.1e-275;
Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;

QY 1 FNLDTENAMTFQFNARGFGQSVVQLQSGRVVVCAPQEIIVAAVNRGSLYOCYSTGSCPT 60
Dd 17 FNLDVEPTTFQSDAGFGQSVVQFGSRLVVGAPLEVAANTGRLYDCAATGACQPI 76
QY 61 RLQVPVEAVNMSGLSLAATSPQLACQPTVHQCSTENTYVKGICFLFGSNLRQPOPK 120
Dd 77 PLHIREAVNNSGLTFLAASGTNGSRLLACGPTLHRVCGNSYSKSGCLLIGSRW-BIIQT 135
QY 121 FPALRGCPQEDDIAFLIDGSGIIPHFRPKRFVSTWMEQLKSKTILPSLMQYSEEF 180
Dd 136 VPDATPECPHEDIVFLIDGSGIDQNDQNMQRGVQAVMGQFGEDTFLFALMQISNLL 195
QY 181 RIHFTKFEQFNPNPSLKPITQLLGRTHATGVRKVRIRRELLNITNGARKNAKILIVI 240
Dd 196 KIHTFTQFTSPSQSLVDPIVQLKLTFTATGILTVTVTQLPHKNGARKSAKILIVI 255
QY 241 TDGEKFGDPIGYEDVPEADREGVIRVIGVADAFPSKRSRQINTIASKPPEDHVPQYN 300
Dd 256 TDGQKYKDPLEYSVPIQAEKAGIIRYAIAGVGHAFQGTARQBLNTISSAPPQDHPKVD 315
QY 301 NFELAKTIQNLREKIPAEIGTGTGSSSSSEHEMSQEGFSAATISNGPLLSTVGSDWAG 360

Db 316 NFRAALGIQLOQEKIYAVEGTQSRASSFQHEMSQEGFTALTNDGLFAGVUSFWSG 375
QY 361 GVELYTSKESKSTFINMTRVDSMDNDAYGAAAIILNRNVQSGLVAGAPRYOHIGLVAMFR 420
Dd 376 GAFLYPNMSPPTFINMSQENVDMEDSYLGYSTELALWKGQNLVLGAPRYOHTKAVIEFT 435
QY 421 QNTGMESNANVKTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETQREGGQSVSCPL 480
Dd 436 QVSRQWRKKAERVTOIGSYFGASLCSVDVDSNGSTDLVLIGAPHYETQREGGQSVSCPL 495
QY 481 PRQQRARWQCDVLYGQGPQWGRFGAALTIVLGDVNGDKLTDVAIGAPEGEDRNGAVYLF 540
Dd 496 PRQQRVQWQCDVLYRGEGHPWGRFGAALTIVLGDVNGDKLTDVAIGAPEGEDRNGAVYLF 555
QY 541 HGTSGSGISPSHSORIASGLSPLOYEQSGSLGGQDLTMDGLVLTWGAOHVILLERSQ 600
Dd 556 HGASEGSGISPSHSORIASGLSPLOYEQSGSLGGQDLTMDGLVLTWGAOHVILLERSQ 615
QY 601 PVLRVKAIMEENPREVAENFECDNDQVWKGAEGERVCLHVQKSTRDLREGOIQSVVT 660
Dd 616 PVLRVKAIMEENPREVAENFECDNDQVWKGAEGERVCLHVQKSTRDLREGOIQSVVT 673
QY 661 YDLALDGRPHSRVAVENETKNSRRQTVGLGTOTCTETLKLQPCNCHDPSPIVLRNF 720
Dd 674 YDLALDGRPHSRVAVENETKNSRRQTVGLGTOTCTETLKLQPCNCHDPSPIVLRNF 733
QY 721 SLVGTPLSAFGLNAPVLAEDAQRLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Dd 734 SLVREPIFSPQNLAPVLAEGSDLPFTASLPPEKNCNDNI CODDLSITFSFMSLDCLVVG 793
QY 781 GPRENVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Dd 794 SSLELVNVTVMNAGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 852
QY 841 SGALKSTCSNHPFPENSEVTFNITDVSLSKSLGNKLLKXAVTSENMMPTNKTEF 900
Dd 853 EG-LSSSCSNHPFPENSEVTFNITDVSLSKSLGNKLLKXAVTSENMMPTNKTEF 911
QY 901 QLELPVKVAVYVTVSHGVSTKYNLF-TASENSTRVMQVQVSNLQSRLEPISLVLPV 959
Dd 912 QLELPVKVAVYVTVSHGVSTKYNLF-TASENSTRVMQVQVSNLQSRLEPISLVLPV 971
QY 960 VRLNQTIVWDRPQVTFSENLSSTCTKRELPSHSDPAELRKAPVNVCSIAVCORICDI 1019
Dd 972 VLLNGVAVVDMVMEAPSQSL--PCVSEKPPQHSDFLTQISKSPMLDCCSIADCLQRCDV 1029
QY 1020 PFGIOEEFNATLKGNLGFDWYIKTSHNHLIVSTAELFNDVSFTLLPFGQAPVRSQTE 1079
Dd 1030 PFSFVQEEELDPTLKNLSFGWVRETLQKVLVWVAEITFDTSVYSQLPQGAEPMAQWE 1089
QY 1080 TKVEPPEVNPPLIVGSSVGGILLALLITALIYALYKLGFPKQYKDMSE 1128
Dd 1090 MVLEBDEVTVNAIPIIMGSSVGALELLALITATLYKLGFPKQYKDMSE 1138

RESULT 15
AAW57491
ID AAW57491 standard; protein; 1161 AA.
XX
XX AAW57491;
XX
XX 24-AUG-1998 (first entry)
XX
XX Human beta2 integrin alpha subunit (alpha d) polypeptide.
XX
XX Beta 2 integrin alpha subunit; alpha d; human; treatment; diabetes;
KW reporter-transactivator construct; arteriosclerosis; atherosclerosis;
KW inflammatory bowel disease; arthritis; multiple sclerosis.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers

```

FT Peptide 1. .16
FT Protein /note= "signal peptide"
FT Domain 17. .1161
FT Domain /note= "mature protein"
FT Domain 17. .1108
FT Domain /note= "extracellular domain"
FT Domain 150. .352
FT Region /note= "i (insertion) domain"
FT Domain 1109. .1128
FT Domain /note= "transmembrane region"
FT Domain 1129. .1161
FT Domain /note= "cytoplasmic domain"
XX US5766850-A.
PN 16-JUN-1998.
XX 21-DEC-1994; 94US-00362652.
XX 23-DEC-1993; 93US-00173497.
PR 05-AUG-1994; 94US-00286889.
XX (ICOS-) ICOS CORP.
XX Van Der Vieren M, Gallatin WM;
XX WPI; 1998-361678/31.
DR N-PSDB; AAV31540.
XX Isolation of DNA encoding protein that binds to integrin subunit - using
XX recombinant cells containing reporter-transactivator construct.
PS Example 5; Col 45-54; 86pp; English.
XX This represents a human beta2 integrin alpha subunit (alpha d)
CC polypeptide. This is used in the methods of the invention for isolating a
CC polynucleotide encoding a protein that binds to alpha d. The method
CC comprises transforming or transfecting host cells with a DNA construct
CC comprising a reporter gene under the control of a promoter regulated by a
CC transcription factor having a DNA-binding domain and an activating
CC domain. A first hybrid DNA sequence encoding a fusion of at least part of
CC alpha d and either the DNA-binding domain or the activating domain of the
CC transcription factor and a library of hybrid DNA sequences encoding
CC fusions of at least part of putative alpha d-binding proteins and the DNA
CC -binding domain or the activating domain of the transcription factor
CC which is not incorporated in the first fusion are expressed in the host
CC cells. The binding of an alpha d-binding protein to alpha d in a
CC particular host cell is detected by determining production of the
CC reporter gene product in the cell. The hybrid DNA sequence encoding the
CC alpha d-binding protein can be isolated from the cell. Alpha d may be
CC useful for treating graft arteriosclerosis, atherosclerosis, diabetes,
CC inflammatory bowel disease, arthritis and multiple sclerosis
XX SQ Sequence 1161 AA;
Query Match 58.2%; Score 3417; DB 2; Length 1161;
Best Local Similarity 59.8%; Pred. No. 2.1e-275;
Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;
QY 1 FNLDENAMTFQENARFGQSVVLOGSRVWVCAPOEIVAAORGSLYOCYSTGSCPEI 60
DB 17 FNLDVEEPTIFQSDAGFGQSVVQFGSLRWVGAPELVAAKOTGRLYDCAATGMCQFI 76
QY 61 RLQVPVAVNMSLGLSLAANTSPFOLLACGPTVHOTCSERNYVKGLCFLFGSNLRQPOK 120
DB 77 PLHIREAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSGKSCLLGSRW-EIIQT 135
QY 121 FPALRGCPQEDSDIAFLIDGSGIIPHPRRMKEFVSTVMEQLKSKTLFSLMOYSEEP 180
DB 136 VPDATPECPHQEMDIVFLIDGSGSIDQDNFNQMKGFVQAVMGQFEGDTLFLALMOYSNLL 195
QY 181 RIHFTKEBFONNPRSLVKPIQOLLGRHTATGAVKAVIRELLNITNGARKNAFKILLI 240

Db 196 KIHFTTQFTSPSQSLVDPIVOLKGLFTTATGILTVVTLFHHKNGARKSAKKILIVI 255
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQBELNTIASPPRDHVFQVN 300
Db 256 TDGQKYKDPLEYSDVIPAOKAGIIRYALGVGHAFQGPATARQBELNTISSAPPQDHVFKVD 315
QY 301 NFEALKTIQNLREKIFAIEGTGTGSSSSPEHMSOBGFSAAITSTNGPLSTVGSYDWAG 360
Db 316 NFAALGSIQKOLQEKIYAVEGTOSRASSSQHMSQEGFTALTMGDLFGAVOSFWSG 375
QY 361 GVFLYTSKEKSTTNNTRVDSMDNDAYLGYAAAIILNRRVQSLVGLGAPRYOHIGLVAMER 420
Db 376 GAFLYPNMGPTTINMSQENVMDRDSYLGYSTELALWKGVQNLVGLGAPRYOHTKAVIFT 435
QY 421 QNTGMESNANVGTQIGAFGASLCSVDVDSNGSTDLVLIGAPHYTYEOTRGQGVSCPL 480
Db 436 QVSRQMKKAEVTTGTQIGSYFGASLCSVDVDSNGSTDLILIGAPHYTYEOTRGQGVSCPL 495
QY 481 PRGORARWQCDVLYGRQGPWGRFGAALTVLGDVNGDKLTDVAIGAPEGEDNRGAYLYF 540
Db 496 PRGORVQWQCDVLYGRQGPWGRFGAALTVLGDVNGDKLTDVAIGAPEGEDNRGAYLYF 555
QY 541 HGTSGSGISPSHSORIASGLSPRLQYFGQSLSGQDLTMDGLVDLTVGAGHVLILRSQ 600
Db 556 HGASEGISPSHSORIASGLSPRLQYFGQSLSGQDLTMDGLVDLTVGAGHVLILRSQ 615
QY 601 PVLRVKAIMBPNPREVARNVFPCNDQVVKGEAGEVVRVCLHVQKSTRDRLRREGQIQSVVT 660
Db 616 PVLKGVAMRFPSPVEVAKAVYRCWEKPSALEAGDAIVCLTIQKESLDQL--GFIQSSVR 673
QY 661 YDLALDSGRPHSRAVFNETKXSTRROTVLGLTQCTETLKLQLPNCIEDPVSPVILNPF 720
Db 674 FDLALDPLRLTSRAIFNETKNPTLRRKTLGLGHCETLKLKLLPDCVEDVYSPILHLNF 733
QY 721 SLVCTPISAFONRPLVLAEDAORLFTALPPEKNCNDNICODDLSTIFSPMSLDCLVVG 780
Db 734 SLVREPISPONRLPVLAVSQDLFTASLPPEKNCQDGLCEGDLGVLTSFSGQLTVG 793
QY 781 GPRFNFVTVVRNDGDSYRTQVTFPFDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 794 SSELNIVTVVWVNADESIGTVVSLYFAGLSHRVSGAQKQPHQSALRLACETV-PTED 852
QY 841 SGALKSTSCSINHDIIPPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 853 EG-LRSSRCSVNHPIFHEGSGTFTVFDVSKATLGDRLMLRASASSENKASSSKATP 911
QY 901 QLELPVYAVVWVTVSHGVSTKYLNPF-TASENTSRVMOHQYQVSNLQORSPLSLVFLVP 959
Db 912 QLELPVYAVVWVTVSHGVSTKYLNPF-TASENTSRVMOHQYQVSNLQORSPLSLVFLVP 971
QY 960 VRLMOTVWDRPQVTFSENLSSTCHTERLPSHSDFLAELKAPVWVNCIAVCORIOCDI 1019
Db 972 VLLNGVAVWVWVMEAPSQSL--PCVSRKRPQHSDFLTQISRSPLMDCSIADCLQFRCDV 1029
QY 1020 PFGIQEENFATLKNLSFDWYIKTSHNHLIVTASILFNDVSVFTLLPGQCAFVRSQTE 1079
Db 1030 PFSVQEBLDTFLKNLSFGVWRETLOKVLVSVABITFDTSVYSQLPGQCAFVRSQTE 1089
QY 1080 TKVEFFEPNPLPIVSGVGLLILALITAAALYKLGFFKRYKDMSE 1128
Db 1090 MVLEDEYVNAIPIMGSSVGLLILALITATLYKLGFFKRYKDMSE 1138

Search completed: June 7, 2004, 17:11:58
Job time : 53.4608 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:16:51 ; Search time 39.8905 Seconds
(without alignments)
8225.189 Million cell updates/sec

Title: US-09-902-481B-6

Perfect score: 5875

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FRQYKQMSGCGPGAEPPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	5875	100.0	1137	10	US-09-902-481A-6
2	5868	99.9	1153	9	US-09-350-259-3
3	5868	99.9	1153	10	US-09-902-481A-1
4	5868	99.9	1153	10	US-09-891-943-3
5	5868	99.9	1153	14	US-10-144-259-30
6	5868	99.9	1153	14	US-10-207-655-176
7	5855	99.7	1137	10	US-09-902-481A-5
8	5852	99.6	1152	9	US-09-945-265-4
9	5845	99.5	1137	10	US-09-902-481A-4
10	5832	99.3	1137	10	US-09-902-481A-3
11	3473	59.1	1163	15	US-10-116-275-204
12	3450	58.7	1163	9	US-09-350-259-4
13	3450	58.7	1163	10	US-09-891-943-4
14	3417	58.2	1161	9	US-09-350-259-2
15	3417	58.2	1161	10	US-09-891-943-2

Sequence 99, Appl
Sequence 99, Appl
Sequence 55, Appl
Sequence 55, Appl
Sequence 53, Appl
Sequence 53, Appl
Sequence 37, Appl
Sequence 37, Appl
Sequence 46, Appl
Sequence 46, Appl
Sequence 1212, Ap
Sequence 2, Appli
Sequence 1, Appli
Sequence 295, App
Sequence 1871, Ap
Sequence 1871, Ap
Sequence 103, App
Sequence 103, App
Sequence 101, App
Sequence 2, Appli
Sequence 2, Appli
Sequence 103, App
Sequence 103, App
Sequence 230, App
Sequence 810, App
Sequence 338, App
Sequence 35, Appl
Sequence 35, Appl
Sequence 4, Appli
Sequence 2, Appli

16 3401.5 57.9 1161 9 US-09-350-259-99
17 3401.5 57.9 1161 10 US-09-891-943-99
18 3239.5 55.1 1161 9 US-09-350-259-55
19 3239.5 55.1 1161 10 US-09-891-943-55
20 3228 54.9 1161 9 US-09-350-259-53
21 3228 54.9 1161 10 US-09-891-943-53
22 3218.5 54.8 1151 9 US-09-350-259-37
23 3218.5 54.8 1151 10 US-09-891-943-37
24 3207 54.6 1155 9 US-09-350-259-46
25 3207 54.6 1155 10 US-09-891-943-46
26 1848 31.5 369 12 US-10-087-192-1212
27 1552.5 26.4 1170 9 US-09-945-265-2
28 1548.5 26.4 1170 12 US-10-361-164-1
29 1525 26.0 1223 16 US-10-408-765A-295
30 1350.5 23.0 1086 16 US-10-408-765A-1871
31 1229.5 20.9 494 9 US-09-350-259-103
32 1229.5 20.9 494 10 US-09-891-943-103
33 1157.5 19.7 413 9 US-09-350-259-101
34 1157.5 19.7 413 10 US-09-891-943-101
35 1149 19.6 1179 14 US-10-177-550-2
36 1149 19.6 1179 15 US-10-173-551-2
37 1102.5 18.8 1151 10 US-09-894-130-103
38 1102.5 18.8 1151 10 US-09-836-353A-103
39 1102.5 18.8 1179 12 US-09-918-715-250
40 1101 18.7 1188 15 US-10-291-265-810
41 1097 18.7 1188 15 US-10-291-265-338
42 1093.5 18.6 1189 10 US-09-984-130-35
43 1093.5 18.6 1189 10 US-09-836-353A-35
44 1093.5 18.6 1189 12 US-10-262-839-4
45 1079 18.4 589 12 US-10-261-164-2

ALIGNMENTS

RESULT 1

US-09-902-481A-6
; Sequence 6, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RPT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902.481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-6

Query Match 100.0%; Score 5875; DB 10; Length 1137;
Best local Similarity 100.0%; Pred. No. 0;
Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGRVVGAPCEIVAAANORGLVOCDYSTGSCPT 60
DB 1 FNLDTENAMTFQENARGFGQSVVQLQGRVVGAPCEIVAAANORGLVOCDYSTGSCPT 60
QY 61 RLQVPEAVNMSLGLSLAATTPPOLLACGPTVHTCSENTYVKGILL FLFGNLRQPOPK 120
DB 61 RLQVPEAVNMSLGLSLAATTPPOLLACGPTVHTCSENTYVKGILL FLFGNLRQPOPK 120
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPDFFRKNKEFVSTVMEQLKKK KTLFSLMQISEEP 180

Db 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKKEFVSTVMEQKSKTLFSLMOYSBEF 180
QY 181 RHFTFKBEONNPNPSLVKPIQLLGRTHATGVRKVIRELLNITNGARKNAFKILLIYI 240
Db 181 RHFTFKBEONNPNPSLVKPIQLLGRTHATGVRKVIRELLNITNGARKNAFKILLIYI 240
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFSEKSRQELNTIASKPRDHVFN 300
Db 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFSEKSRQELNTIASKPRDHVFN 300
QY 301 NFPAKTIQNLREKIFAEGTGTSSSFEHEMSQEGFSAATTSNGPLLSVTGSDWAG 360
Db 301 NFPAKTIQNLREKIFAEGTGTSSSFEHEMSQEGFSAATTSNGPLLSVTGSDWAG 360
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAALILNRRVQSLVLCAPRYOHIGLVAMER 420
Db 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAALILNRRVQSLVLCAPRYOHIGLVAMER 420
QY 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSVCP 480
Db 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSVCP 480
QY 481 PRQORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVILF 540
Db 481 PRQORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVILF 540
QY 541 HGTSGGSISSHQRIAGSKLSRLOVFGQSLGGQDLTMDGLVDTLVGAGHVLRLSQ 600
Db 541 HGTSGGSISSHQRIAGSKLSRLOVFGQSLGGQDLTMDGLVDTLVGAGHVLRLSQ 600
QY 601 PVLVRKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLRREGQIOSVVT 660
Db 601 PVLVRKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLRREGQIOSVVT 660
QY 661 YDLALDSGRPHSRAVFNTRRQTVLGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720
Db 661 YDLALDSGRPHSRAVFNTRRQTVLGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720
QY 721 SLVGTPLSAGNLRPLVLAEDAORLFTALPFPEKNCNDNICODDLSITFPMSLDCLVVG 780
Db 721 SLVGTPLSAGNLRPLVLAEDAORLFTALPFPEKNCNDNICODDLSITFPMSLDCLVVG 780
QY 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLVYKVTSLQNRQSRSWRLACASSTEV 840
Db 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLVYKVTSLQNRQSRSWRLACASSTEV 840
QY 841 SGALKSTSINHPDIPFENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 841 SGALKSTSINHPDIPFENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
QY 901 QLELPVKYAVVWVTSKGVSTKYLNTASNTSRVMOHQYQVSNLQORSLSPLSLVFLVPV 960
Db 901 QLELPVKYAVVWVTSKGVSTKYLNTASNTSRVMOHQYQVSNLQORSLSPLSLVFLVPV 960
QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDPLAELRKA PVNCSIAVCORIQCIP 1020
Db 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDPLAELRKA PVNCSIAVCORIQCIP 1020
QY 1021 FFGIQEERFNATLKNLSFDWIKTSHNHLITVSTABILFNDSVFTLLPGOGAFVRSQTE 1080
Db 1021 FFGIQEERFNATLKNLSFDWIKTSHNHLITVSTABILFNDSVFTLLPGOGAFVRSQTE 1080
QY 1081 KVEPFEVNPPLIVGSSVGLLALITAALYKLGFFKQYKDMSEGGPPGAEPPQ 1137
Db 1081 KVEPFEVNPPLIVGSSVGLLALITAALYKLGFFKQYKDMSEGGPPGAEPPQ 1137

RESULT 2
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:

; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: NO. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,989
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.9%; Score 5868; DB 9; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLDTENAMTPOENARGFGQSVQVQGGSRVVVVGAPQETVAANQORSLYOCDYSTGSCPEI 60
DB 17 FNLDTENAMTPOENARGFGQSVVQVQGGSRVVVVGAPQETVAANQORSLYOCDYSTGSCPEI 76
QY 61 RLQVPEAVNMSLGLSLAATSPOLLACGPTVHTCSENTYVVKGLCLFPGNLRQPOPK 120
DB 77 RLQVPEAVNMSLGLSLAATSPOLLACGPTVHTCSENTYVVKGLCLFPGNLRQPOPK 136
QY 121 PFEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKKEFVSTVMEQKSKTLFSLMOYSBEF 180
DB 137 PFEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKKEFVSTVMEQKSKTLFSLMOYSBEF 196
QY 181 RHFTFKBEONNPNPSLVKPIQLLGRTHATGVRKVIRELLNITNGARKNAFKILLIYI 240
DB 197 RHFTFKBEONNPNPSLVKPIQLLGRTHATGVRKVIRELLNITNGARKNAFKILLIYI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFSEKSRQELNTIASKPRDHVFN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFSEKSRQELNTIASKPRDHVFN 316
QY 301 NFPAKTIQNLREKIFAEGTGTSSSFEHEMSQEGFSAATTSNGPLLSVTGSDWAG 360
DB 317 NFPAKTIQNLREKIFAEGTGTSSSFEHEMSQEGFSAATTSNGPLLSVTGSDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAALILNRRVQSLVLCAPRYOHIGLVAMER 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAALILNRRVQSLVLCAPRYOHIGLVAMER 436
QY 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSVCP 480
DB 437 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSVCP 496
QY 481 PRQORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVILF 540
DB 497 PRQORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVILF 556
QY 541 HGTSGGSISSHQRIAGSKLSRLOVFGQSLGGQDLTMDGLVDTLVGAGHVLRLSQ 600
DB 557 HGTSGGSISSHQRIAGSKLSRLOVFGQSLGGQDLTMDGLVDTLVGAGHVLRLSQ 616
QY 601 PVLVRKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLRREGQIOSVVT 660
DB 617 PVLVRKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLRREGQIOSVVT 676
QY 661 YDLALDSGRPHSRAVFNTRRQTVLGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720

Db 677 YDLALDSGRPHSRVAFNETKNSRQVGLGTQTCETLKLQLPNCIEDPVPVLRNF 736
QY 721 SLVGTPLSAFGLNLRVLAEDAQRFTALFPFKXKCGNDNICODDLSITFSFMSLDCLVWG 780
Db 737 SLVGTPLSAFGLNLRVLAEDAQRFTALFPFKXKCGNDNICODDLSITFSFMSLDCLVWG 796
QY 781 GPREFNVTVVRNDGEDSYRQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKANVTSENNMPRTNKTEP 900
Db 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKANVTSENNMPRTNKTEP 916
QY 901 QLELPVKYAVTMVTVSHGVSTKYLNFASENTSRVMQHOYQVSNLQGRSLPISLVFLVPV 960
Db 917 QLELPVKYAVTMVTVSHGVSTKYLNFASENTSRVMQHOYQVSNLQGRSLPISLVFLVPV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCISIAVCORIQQDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCISIAVCORIQQDIP 1036
QY 1021 PFGIOEBFNATLKGNSLSDWIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1080
Db 1037 PFGIOEBFNATLKGNSLSDWIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1096
QY 1081 KVEPEVEPNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1137
Db 1097 KVEPEVEPNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1153

RESULT 3

US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimooka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RNK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (17)..()
; OTHER INFORMATION:
US-09-902-481A-1

Query Match 99.9%; Score 5868; DB 10; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFQENARGFCQSVVQLQGSRRVVVGAQPEIIVAAHQSGSLYQCOYSTGSCBPI 60
Db 17 FNLDTENAMTFQENARGFCQSVVQLQGSRRVVVGAQPEIIVAAHQSGSLYQCOYSTGSCBPI 76
QY 61 RLQVPEAVNNSIGLSLAATTPPOLLACGPVTHQTCSENTVYKGLCFPLGNSLRQCPQK 120
Db 77 RLQVPEAVNNSIGLSLAATTPPOLLACGPVTHQTCSENTVYKGLCFPLGNSLRQCPQK 136
QY 121 FFEALRGCPQEDSDIAFLDGSIIIPHDPRFMKEFVSTWMEQIKKSKTLPSLMQYSEEF 180

Db 137 FFEALRGCPQEDSDIAFLDGSIIIPHDPRFMKEFVSTWMEQIKKSKTLPSLMQYSEEF 196
QY 181 RIHFTFKFQNNPNPSLVKPIIQTLLGRTHATGVVRKVTRELLNITNGARKNAFKILIVI 240
Db 197 RIHFTFKFQNNPNPSLVKPIIQTLLGRTHATGIRKVRRELFNITNGARKNAFKILIVI 256
QY 241 TDEKTECDPLGYEDVTPEADREGVIRVIGVGDAFTSEKROELNITIASPPRDHVFQVN 300
Db 257 TDEKTECDPLGYEDVTPEADREGVIRVIGVGDAFTSEKROELNITIASPPRDHVFQVN 316
QY 301 NFBAKTIQNLAREKIPAIETGTGSSSSPEHMSQEGFSAATSNGLPLSTVGSYDWAG 360
Db 317 NFBAKTIQNLAREKIPAIETGTGSSSSPEHMSQEGFSAATSNGLPLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTINMTRVDSDAVDAYLGAARAILLNRVQSLVLAGAPRYCHIGLVAMER 420
Db 377 GVFLYTSKEKSTINMTRVDSDAVDAYLGAARAILLNRVQSLVLAGAPRYCHIGLVAMER 436
QY 421 QNTGWMESNANVKTQICAYFGASLSCVDVDSNGSTDVLVIGAPHYEYQTRGGQVSCPL 480
Db 437 QNTGWMESNANVKTQICAYFGASLSCVDVDSNGSTDVLVIGAPHYEYQTRGGQVSCPL 496
QY 481 PRGQARWQCDAVLYGEGQGFWRPGALTVLDVNGDKLTDVAI GAPGEDNREGAVYLP 540
Db 497 PRGQARWQCDAVLYGEGQGFWRPGALTVLDVNGDKLTDVAI GAPGEDNREGAVYLP 556
QY 541 HGTSGSGISPSHSORIASGKLSPLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSORIASGKLSPLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIMBFPNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTDRRLREGQIQSVVT 660
Db 617 PVLRVKAIMBFPNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTDRRLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSRQVGLGTQTCETLKLQLPNCIEDPVPVLRNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSRQVGLGTQTCETLKLQLPNCIEDPVPVLRNF 736
QY 721 SLVGTPLSAFGLNLRVLAEDAQRFTALFPFKXKCGNDNICODDLSITFSFMSLDCLVWG 780
Db 737 SLVGTPLSAFGLNLRVLAEDAQRFTALFPFKXKCGNDNICODDLSITFSFMSLDCLVWG 796
QY 781 GPREFNVTVVRNDGEDSYRQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKANVTSENNMPRTNKTEP 900
Db 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKANVTSENNMPRTNKTEP 916
QY 901 QLELPVKYAVTMVTVSHGVSTKYLNFASENTSRVMQHOYQVSNLQGRSLPISLVFLVPV 960
Db 917 QLELPVKYAVTMVTVSHGVSTKYLNFASENTSRVMQHOYQVSNLQGRSLPISLVFLVPV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCISIAVCORIQQDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCISIAVCORIQQDIP 1036
QY 1021 PFGIOEBFNATLKGNSLSDWIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1080
Db 1037 PFGIOEBFNATLKGNSLSDWIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1096
QY 1081 KVEPEVEPNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1137
Db 1097 KVEPEVEPNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1153

RESULT 4
US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.


```

; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-891-943-3

Query Match          99.9%; Score 5868; DB 10; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCBPI 76

QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGICFLFGSNLRQOQPK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGICFLFGSNLRQOQPK 136

QY 121 FPEALGCGQEDSDIAFLIDGSGSIIPHDFRMRKPEFVSTMQLKSKTFLSLMOYSEBF 180
DB 137 FPEALGCGQEDSDIAFLIDGSGSIIPHDFRMRKPEFVSTMQLKSKTFLSLMOYSEBF 196

QY 181 RHFTFKPONNPNRSLVKPITOLLGRTHGTATGKVRKRELLNITNGARKNAFKLLIYI 240
DB 197 RHFTFKPONNPNRSLVKPITOLLGRTHGTATGKVRKRELLNITNGARKNAFKLLIYI 256

QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQBLNTIASKPRDHVFQVN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQBLNTIASKPRDHVFQVN 316

QY 301 NFPAKTIQNLREKIFAEIGTGTGSSSEFHEMSQEGSAATSNGLLSTVGSYDMAG 360
DB 317 NFPAKTIQNLREKIFAEIGTGTGSSSEFHEMSQEGSAATSNGLLSTVGSYDMAG 376

QY 361 GVFLYTSKEKSTFNTRVDSMDNDAYLGAAAIILNRRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFNTRVDSMDNDAYLGAAAIILNRRVQSLVGLGAPRYQHIGLVAMFR 436

QY 421 QNTGMESNANVKGITQIGAFGASLCSVDVDSNGSTDVLIGAPHYEQTRGQGVSCPL 480
DB 437 QNTGMESNANVKGITQIGAFGASLCSVDVDSNGSTDVLIGAPHYEQTRGQGVSCPL 496

QY 481 PRGORARWQCDVLYGSGQOPWGFPGAAALVLDGVDNGDKLTDVAIGAPGEDNRGAVILF 540
DB 497 PRGORARWQCDVLYGSGQOPWGFPGAAALVLDGVDNGDKLTDVAIGAPGEDNRGAVILF 556

QY 541 HGTSGSGISPSHSGRIAGSKLSRPLQYFGQSLSGQDPLWDGLVLDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSGRIAGSKLSRPLQYFGQSLSGQDPLWDGLVLDLTVGAQGHVLLRSQ 616

QY 601 PVLRVKALMFPNPREVARNVPECDNVKKEAGEVRVCHVOKSTRDLRREGQIOSVVT 660
DB 617 PVLRVKALMFPNPREVARNVPECDNVKKEAGEVRVCHVOKSTRDLRREGQIOSVVT 676

QY 661 YDLALDSGRPHSRVAFNETKNSRQOVVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
DB 677 YDLALDSGRPHSRVAFNETKNSRQOVVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 736

721 SLVGTPLSAFNGNLRPVLAEADAQRLFTALFPFEKNCNGNDNICQDDLSITFSFMSLDCLVVG 780
737 SLVGTPLSAFNGNLRPVLAEADAQRLFTALFPFEKNCNGNDNICQDDLSITFSFMSLDCLVVG 796

781 GPREFNVTVVRNDGSDSYTOTTFPEPLDI SVKXKUSTLONORSORSWRLACBSASSTEV 840
797 GPREFNVTVVRNDGSDSYTOTTFPEPLDI SVKXKUSTLONORSORSWRLACBSASSTEV 856

841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANKVTSENNMPRTNKTPE 900
857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANKVTSENNMPRTNKTPE 916

901 QLELPVKYAVTMVTVSHGVSTKYLNFNTASENTSRVMOHQVSNLQORSLSISLVLVLPV 960
917 QLELPVKYAVTMVTVSHGVSTKYLNFNTASENTSRVMOHQVSNLQORSLSISLVLVLPV 976

961 RLNCTVIWDRPOVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVVCNCSIAVCQRIQCDIP 1020
977 RLNCTVIWDRPOVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVVCNCSIAVCQRIQCDIP 1036

1021 PFGIOEBFNATLKNLSFDWYIKTSHRHLLIVSTABILFNDVSVELLPGOCAPVRSOTET 1080
1037 PFGIOEBFNATLKNLSFDWYIKTSHRHLLIVSTABILFNDVSVELLPGOCAPVRSOTET 1096

1081 KVEPPEVNPPLIIVGSSVGGLLILLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1137
1097 KVEPPEVNPPLIIVGSSVGGLLILLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1153

RESULT 5
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnsaut, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-144-259-30

Query Match          99.9%; Score 5868; DB 14; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCBPI 76

QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGICFLFGSNLRQOQPK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGICFLFGSNLRQOQPK 136

QY 121 FPEALGCGQEDSDIAFLIDGSGSIIPHDFRMRKPEFVSTMQLKSKTFLSLMOYSEBF 180
DB 137 FPEALGCGQEDSDIAFLIDGSGSIIPHDFRMRKPEFVSTMQLKSKTFLSLMOYSEBF 196

QY 181 RHFTFKPONNPNRSLVKPITOLLGRTHGTATGKVRKRELLNITNGARKNAFKLLIYI 240
DB 197 RHFTFKPONNPNRSLVKPITOLLGRTHGTATGKVRKRELLNITNGARKNAFKLLIYI 256
```

```
QY 241 TDGEKFGDPLGYEDVIVPEADREGVIRVYGVGDAPRSEKSRQELNTIASKPRDHVFOVN 300
DB 257 TDGEKFGDPLGYEDVIVPEADREGVIRVYGVGDAPRSEKSRQELNTIASKPRDHVFOVN 316
QY 301 NFEALKTIONQUREKI FAIEGTQTGSSSSFFHEMSQEGFSAITNSGSLTGVSYDWAG 360
DB 317 NFEALKTIONQUREKI FAIEGTQTGSSSSFFHEMSQEGFSAITNSGSLTGVSYDWAG 376
QY 361 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 436
QY 421 QNTGWSNANVKGQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQVSVCP 480
DB 437 QNTGWSNANVKGQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQVSVCP 496
QY 481 PRGQARWQCDVAVLYGEOGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVYLF 540
DB 497 PRGQARWQCDVAVLYGEOGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLTVGAQHVLILRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLTVGAQHVLILRSQ 616
QY 601 PVLRVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 660
DB 617 PVLRVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRAVFNETKNSRTRQTQVLGTQTCETLKLQPCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPHSRAVFNETKNSRTRQTQVLGTQTCETLKLQPCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAFGMLRPVLAEDAORLFTALPPFKKNCNDNICODDLSITFSFMSLCLVVG 780
DB 737 SLVGTPLSAFGMLRPVLAEDAORLFTALPPFKKNCNDNICODDLSITFSFMSLCLVVG 796
QY 781 GPREFNVTVVRNDGEDSVRTQVTQVTFPPFLDLSTRKYSTLQNSQSRWFLACESASSTEV 840
DB 797 GPREFNVTVVRNDGEDSVRTQVTQVTFPPFLDLSTRKYSTLQNSQSRWFLACESASSTEV 856
QY 841 SGALKSTCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEP 900
DB 857 SGALKSTCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEP 916
QY 901 QLELPVKYAVVWVTSKYSTKYNFTASNTSRVMQHOYQVSNLQORSPLISLVLV 960
DB 917 QLELPVKYAVVWVTSKYSTKYNFTASNTSRVMQHOYQVSNLQORSPLISLVLV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVMNCSTAVCORIQCDIP 1020
DB 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVMNCSTAVCORIQCDIP 1036
QY 1021 FFGIQEENATLKENLSFOWYIKTSENHLLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1080
DB 1037 FFGIQEENATLKENLSFOWYIKTSENHLLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1096
QY 1081 KVEPFFVNPPLPIVGVSSVGLLLALLITAAALYKLGFFKQKDMMSBGGPGAEPQ 1137
DB 1097 KVEPFFVNPPLPIVGVSSVGLLLALLITAAALYKLGFFKQKDMMSBGGPGAEPQ 1153
```

RESULT 6

```
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
```

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 176

; LENGTH: 1153

; TYPE: PR

; ORGANISM: Homo sapiens

US-10-207-655-176

Query Match 99.9%; Score 5868; DB 14; Length 1153;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1; Indels 0; Gaps 0;

Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDENAMTFQENARFGQSVVOLQSGRVVVGAPQBIIVANQORGSLYQCDYSTGSCPEI 60

DB 17 FNLDENAMTFQENARFGQSVVOLQSGRVVVGAPQBIIVANQORGSLYQCDYSTGSCPEI 76

QY 61 RLQVPEAVANVKSGLSLAAITSPPELLACGPTVHOTCSENTYVKGLCFLFGSNLRQOPQK 120

DB 77 RLQVPEAVANVKSGLSLAAITSPPELLACGPTVHOTCSENTYVKGLCFLFGSNLRQOPQK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRRMEKFVSTVMEQKKSKTLPFLMOYSEEF 180

DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRRMEKFVSTVMEQKKSKTLPFLMOYSEEF 196

QY 181 RIHTFPKQFQNNPRSLVXPIITOLGRTHATGVRVIRELLNITNGARKNAFKILTVI 240

DB 197 RIHTFPKQFQNNPRSLVXPIITOLGRTHATGVRVIRELLNITNGARKNAFKILTVI 256

QY 241 TDGEKFGDPLGYEDVIVPEADREGVIRVYGVGDAPRSEKSRQELNTIASKPRDHVFOVN 300

DB 257 TDGEKFGDPLGYEDVIVPEADREGVIRVYGVGDAPRSEKSRQELNTIASKPRDHVFOVN 316

QY 301 NFEALKTIONQUREKI FAIEGTQTGSSSSFFHEMSQEGFSAITNSGSLTGVSYDWAG 360

DB 317 NFEALKTIONQUREKI FAIEGTQTGSSSSFFHEMSQEGFSAITNSGSLTGVSYDWAG 376

QY 361 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 420

DB 377 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 436

QY 421 QNTGWSNANVKGQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQVSVCP 480

DB 437 QNTGWSNANVKGQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQVSVCP 496

QY 481 PRGQARWQCDVAVLYGEOGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVYLF 540

DB 497 PRGQARWQCDVAVLYGEOGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVYLF 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLTVGAQHVLILRSQ 600

DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLTVGAQHVLILRSQ 616

QY 601 PVLRVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 660

DB 617 PVLRVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 676

QY 661 YDLALDSGRPHSRAVFNETKNSRTRQTQVLGTQTCETLKLQPCIEDPVSPIVLRNF 720

DB 677 YDLALDSGRPHSRAVFNETKNSRTRQTQVLGTQTCETLKLQPCIEDPVSPIVLRNF 736

QY 721 SLVGTPLSAFGMLRPVLAEDAORLFTALPPFKKNCNDNICODDLSITFSFMSLCLVVG 780

DB 737 SLVGTPLSAFGMLRPVLAEDAORLFTALPPFKKNCNDNICODDLSITFSFMSLCLVVG 796

QY 781 GPREFNVTVVRNDGEDSVRTQVTQVTFPPFLDLSTRKYSTLQNSQSRWFLACESASSTEV 840

DB 797 GPREFNVTVVRNDGEDSVRTQVTQVTFPPFLDLSTRKYSTLQNSQSRWFLACESASSTEV 856

QY 841 SGALKSTCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEP 900

DB 857 SGALKSTCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEP 916

QY 901 QLELPVKYAVVWVTSKYSTKYNFTASNTSRVMQHOYQVSNLQORSPLISLVLV 960

Db 917 QLELPVYAYVMVTSVSHGSKYKLTNTASNTSRVVMQHQVQVSNLQORSIPISLVLFPV 976
Qy 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSPSHSDFLAELRKAPVNVNCSIAVCORIQCIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSPSHSDFLAELRKAPVNVNCSIAVCORIQCIP 1036
Qy 1021 FFGIOEFNATLKNLSFDWIKTSNHLIIVSTABILLFNDVSFTLLPGQAFVRSQTET 1080
Db 1037 FFGIOEFNATLKNLSFDWIKTSNHLIIVSTABILLFNDVSFTLLPGQAFVRSQTET 1096
Qy 1081 KVEPPEVNPPLPIVGVSSVGGILLALITAALYKLGFFKQYKDMMSGGPPGAEPQ 1137
Db 1097 KVEPPEVNPPLPIVGVSSVGGILLALITAALYKLGFFKQYKDMMSGGPPGAEPQ 1153

RESULT 7
US-09-902-481A-5
; Sequence 5, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimada, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902.481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-5

Query Match 99.7%; Score 5855; DB 10; Length 1137;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1127; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTQENARFGQSVVOLQGSRRVVGAPQBIIVANQSGSIYQCDYSTGSCPI 60
Db 1 FNLDTENAMTQENARFGQSVVOLQGSRRVVGAPQBIIVANQSGSIYQCDYSTGSCPI 60

Qy 61 RLQVPVAVNMSLGLSLAATSPOLLACGPTVHTCSENTYVKGCLFGLGSLNLRQOPQK 120
Db 61 RLQVPVAVNMSLGLSLAATSPOLLACGPTVHTCSENTYVKGCLFGLGSLNLRQOPQK 120

Qy 121 FPEALRGCPQSDSIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTLFSLMQYSEEP 180
Db 121 FPEALRGCPQSDSIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTLFSLMQYSEEP 180

Qy 181 RIHFTKEFQNNPRLSKVPIQTLGKTHATGVRKVIKRELLTNGARKNAFKILIVI 240
Db 181 RIHFTKEFQNNPRLSKVPIQTLGKTHATGVRKVIKRELLTNGARKNAFKILILI 240

Qy 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDADFSEKSRQELNTIASKPPRDHVFQVN 300
Db 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDADFSEKSRQELNTIASKPPRDHVFQIN 300

Qy 301 NPEALKTIQNLREKIFAIETQGTSSSFHEMSQGFSAITNSGELLSTVGSYDAG 360
Db 301 NPEALKTIQNLREKIFAIETQGTSSSFHEMSQGFSAITNSGELLSTVGSYDAG 360

Qy 361 GVFLYTSKEKSTFINMTVRDSDMDAYLGYAAIILNRVQSLVIGAPRYOHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFINMTVRDSDMDAYLGYAAIILNRVQSLVIGAPRYOHIGLVAMFR 420

Qy 421 QNTGWNESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVUCPL 480

RESULT 8

US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimada, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; FILE REFERENCE: CBN-002CP
; CURRENT APPLICATION NUMBER: US/09/945,265
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,700
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-265-4

Query Match 99.6%; Score 5852.5; DB 9; Length 1152;
Best Local Similarity 99.6%; Pred. No. 0;

Matches 1132; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFQENARGFQSVVQLOGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFQSVVQLOGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVAVNMSLGLSLAATTSPPQILACGPTVHQCSTENTYVKGICFLFGSNLRQOPQK 120
DB 77 RLQVPVAVNMSLGLSLAATTSPPQILACGPTVHQCSTENTYVKGICFLFGSNLRQOPQK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPFRMKBFVSTVMEQKSKTLFSLMQYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPFRMKBFVSTVMEQKSKTLFSLMQYSEEF 196

QY 181 RIHFTFKFQNNPNSRLVXPIIQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240
DB 197 RIHFTFKFQNNPNSRLVXPIIQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 256

QY 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQBELNTIASKPPRDHVPQVN 300
DB 257 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQBELNTIASKPPRDHVPQVN 316

QY 301 NFEALKTIQNLREKIFAIEGTOTGSSSSPEHEMSQEGFSAATISNGPILLSTVGSYDWAQ 360
DB 317 NFEALKTIQNLREKIFAIEGTOTGSSSSPEHEMSQEGFSAATISNGPILLSTVGSYDWAQ 376

QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGVAALIIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGVAALIIILNRVQSLVGLGAPRYQHIGLVAMFR 436

QY 421 QNTGWSNANVKTQIAGYFASLCSVDVDSNGSTDVLVIGAPHYVEOTRGQSVCP 480
DB 437 QNTGWSNANVKTQIAGYFASLCSVDVDSNGSTDVLVIGAPHYVEOTRGQSVCP 496

QY 481 PRGQARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAI GAPBEDNRGAVYLF 540
DB 497 PRGQARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAI GAPBEDNRGAVYLF 555

QY 541 HGTSGSGISPSHSORIASGSKSLPQYFGOSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 600
DB 556 HGTSGSGISPSHSORIASGSKSLPQYFGOSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 615

QY 601 PVLAVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRLEGQIQSVVT 660
DB 616 PVLAVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRLEGQIQSVVT 675

QY 661 YDLALDSGRPHSRVAVNETKSTRQTVGLGTQTCETLKLQPNICIEDPVSPVILRLNF 720
DB 676 YDLALDSGRPHSRVAVNETKSTRQTVGLGTQTCETLKLQPNICIEDPVSPVILRLNF 735

QY 721 SLVGTPLSAFNLRLVLAEDAQRLEFTALFPFEKXCGNDNICQDDLSTIFSMSLDCLVVG 780
DB 736 SLVGTPLSAFNLRLVLAEDAQRLEFTALFPFEKXCGNDNICQDDLSTIFSMSLDCLVVG 795

QY 781 GPREFNVTVVRNDGDSYRTQVTFPPPLDLVSRKYSTLQNRQSRWRLACESASTEV 840
DB 796 GPREFNVTVVRNDGDSYRTQVTFPPPLDLVSRKYSTLQNRQSRWRLACESASTEV 855

QY 841 SGALKSTSCSINRHPFPENSEVFNITFDVDSKASLGNKLLKXANTYSENMPRTNKTEF 900
DB 856 SGALKSTSCSINRHPFPENSEVFNITFDVDSKASLGNKLLKXANTYSENMPRTNKTEF 915

QY 901 QLELPVKYAVTWVTSRGVSKYLNFTASENTSRVMQHOYQVSNLQORSILPISLPLVPV 960
DB 916 QLELPVKYAVTWVTSRGVSKYLNFTASENTSRVMQHOYQVSNLQORSILPISLPLVPV 975

QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDIP 1020
DB 976 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDIP 1035

QY 1021 PFGIQEENFATLKNLSFDVAVIKTSNNHLLIVSTAEILNDSVFTLLRQGGPVSOTET 1080
DB 1036 PFGIQEENFATLKNLSFDVAVIKTSNNHLLIVSTAEILNDSVFTLLRQGGPVSOTET 1095

QY 1081 KVRPFVFNPLPLIVGSSVGLLILLALITAAALYKLGFFKQYKDMSEGPPGABPQ 1137
DB 1096 KVRPFVFNPLPLIVGSSVGLLILLALITAAALYKLGFFKQYKDMSEGPPGABPQ 1152

RESULT 9
US-09-902-481A-4
; Sequence 4, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RTT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-4

Query Match 99.5%; Score 5845; DB 10; Length 1137;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1126; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFQSVVQLOGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
DB 1 FNLDTENAMTFQENARGFQSVVQLOGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60

QY 61 RLQVPVAVNMSLGLSLAATTSPPQILACGPTVHQCSTENTYVKGICFLFGSNLRQOPQK 120
DB 61 RLQVPVAVNMSLGLSLAATTSPPQILACGPTVHQCSTENTYVKGICFLFGSNLRQOPQK 120

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPFRMKBFVSTVMEQKSKTLFSLMQYSEEF 180
DB 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPFRMKBFVSTVMEQKSKTLFSLMQYSEEF 180

QY 181 RIHFTFKFQNNPNSRLVXPIIQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240
DB 181 RIHFTFKFQNNPNSRLVXPIIQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240

QY 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQBELNTIASKPPRDHVPQVN 300
DB 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQBELNTIASKPPRDHVPQVN 300

QY 301 NFEALKTIQNLREKIFAIEGTOTGSSSSPEHEMSQEGFSAATISNGPILLSTVGSYDWAQ 360
DB 301 NFEALKTIQNLREKIFAIEGTOTGSSSSPEHEMSQEGFSAATISNGPILLSTVGSYDWAQ 360

QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGVAALIIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGVAALIIILNRVQSLVGLGAPRYQHIGLVAMFR 420

QY 421 QNTGWSNANVKTQIAGYFASLCSVDVDSNGSTDVLVIGAPHYVEOTRGQSVCP 480
DB 421 QNTGWSNANVKTQIAGYFASLCSVDVDSNGSTDVLVIGAPHYVEOTRGQSVCP 480

QY 481 PRGQARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAI GAPBEDNRGAVYLF 540
DB 481 PRGQARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAI GAPBEDNRGAVYLF 540

QY 541 HGTSGSGISPSHSORIASGSKSLPQYFGOSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 600
DB 541 HGTSGSGISPSHSORIASGSKSLPQYFGOSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 600

Db 541 HGTSGSGLSPSHSQRIAGSKLSPLRQYRGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 600
Qy 601 PVLRYKAIMFNPVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 601 PVLRYKAIMFNPVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Qy 661 YDLALDSGRPHSRVFNKSTRQTOVLGLTTCETLKLQKNCIEDVPSPVLRLNF 720
Db 661 YDLALDSGRPHSRVFNKSTRQTOVLGLTTCETLKLQKNCIEDVPSPVLRLNF 720
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRFTALPFPFKKNCNDNICQDDLSITPFSMSLDCLVVG 780
Db 721 SLVGTPLSAFGLNRPVLAEDAQRFTALPFPFKKNCNDNICQDDLSITPFSMSLDCLVVG 780
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPFLDLSYKYSTLQORSORSWRLACASASTEV 840
Db 781 GPREFNVTVVRNDGEDSYRTQVTFPFLDLSYKYSTLQORSORSWRLACASASTEV 840
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGKLLKANVTSENNMPTNKTEF 900
Db 841 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGKLLKANVTSENNMPTNKTEF 900
Qy 901 QLELPVKYAVYVWVTSKYNLFTASENTSRVMOHQVSNLQORSIPISLVLVPV 960
Db 901 QLELPVKYAVYVWVTSKYNLFTASENTSRVMOHQVSNLQORSIPISLVLVPV 960
Qy 961 RLNQTVIWDREPQVTFSEKLSSTCKTKERLPSHSDFLAELKAPVYVNCIAVCQRIQCDIP 1020
Db 961 RLNQTVIWDREPQVTFSEKLSSTCKTKERLPSHSDFLAELKAPVYVNCIAVCQRIQCDIP 1020
Qy 1021 PFGIOEEFNATLKNLSFDWIKTSHNHLIIVSTAEILFNDSPVTLPGQAFVRSQTET 1080
Db 1021 PFGIOEEFNATLKNLSFDWIKTSHNHLIIVSTAEILFNDSPVTLPGQAFVRSQTET 1080
Qy 1081 KVEPFEVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGABEQ 1137
Db 1081 KVEPFEVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGABEQ 1137

RESULT 10

US-09-902-481a-3
; Sequence 3, Application US/09902481a
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimooka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481a
; PRIOR FILING DATE: 2001-07-09
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481a-3

Query March 99.3%; Score 5832; DB 10; Length 1137;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1122; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTQENARGFGQSVQLOGSRVVGAPQBIIVAAHQSGSLYQCDYSTGSCETP 60
Db 1 FNLDTENAMTQENARGFGQSVQLOGSRVVGAPQBIIVAAHQSGSLYQCDYSTGSCETP 60
Qy 61 RLQVPEAVNNSLGLSLAATTSPPQLLACGPTVHTQCTSENTYVYKGLCFPLGSLNRQOPK 120

RESULT 11

Db 61 RLQVPEAVNNSLGLSLAATTSPPQLLACGPTVHTQCTSENTYVYKGLCFPLGSLNRQOPK 120
Qy 121 PPEALRGCPQSDSDIAPLIDGSGSIIIPDPRMKEFPYSTVMBOLKSKTLPSLMOYSEEF 180
Db 121 PPEALRGCPQSDSDIAPLIDGSGSIIIPDPRMKEFPYSTVMBOLKSKTLPSLMOYSEEF 180
Qy 181 RIHFTFKFQFQNPSPRSUVKIIITOLGRTHATGVRKVIIRLLNITNGARKNAKFIILVI 240
Db 181 RIHFTFKFQFQNPSPRSUVKIIITOLGRTHATGVRKVIIRLLNITNGARKNAKFIILVI 240
Qy 241 TDGSKFGDPLGYEDVIEADREGVIRYVIGVDAPFRSEKSRQELNITIASKPPRDHVPQN 300
Db 241 TDGSKFGDPLGYEDVIEADREGVIRYVIGVDAPFRSEKSRQELNITIASKPPRDHVPQN 300
Qy 301 NPEALKTIQONLRKIIIPALBCTQTSSTSSSEHMSQEGPSAAITSGPILSTVGSYDWAG 360
Db 301 NPEALKTIQONLRKIIIPALBCTQTSSTSSSEHMSQEGPSAAITSGPILSTVGSYDWAG 360
Qy 361 GVFYLYTSKEKSTFTNMTFRVDSMDNDAYLGAAAIILNRRVQSLVLAGAPRYOHIGLVAMPR 420
Db 361 GVFYLYTSKEKSTFTNMTFRVDSMDNDAYLGAAAIILNRRVQSLVLAGAPRYOHIGLVAMPR 420
Qy 421 QNTGMBESNANVKTOIGAYEGASLCSVDVDSNGSDTLVLIQAPHYEOTRGQSVCP 480
Db 421 QNTGMBESNANVKTOIGAYEGASLCSVDVDSNGSDTLVLIQAPHYEOTRGQSVCP 480
Qy 481 PRGQARWQCDVLYYGRQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNREGAVYLF 540
Db 481 PRGQARWQCDVLYYGRQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNREGAVYLF 540
Qy 541 HGTSGSGLSPSHSQRIAGSKLSPLRQYRGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 600
Db 541 HGTSGSGLSPSHSQRIAGSKLSPLRQYRGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 600
Qy 601 PVLRYKAIMFNPVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 601 PVLRYKAIMFNPVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Qy 661 YDLALDSGRPHSRVFNKSTRQTOVLGLTTCETLKLQKNCIEDVPSPVLRLNF 720
Db 661 YDLALDSGRPHSRVFNKSTRQTOVLGLTTCETLKLQKNCIEDVPSPVLRLNF 720
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRFTALPFPFKKNCNDNICQDDLSITPFSMSLDCLVVG 780
Db 721 SLVGTPLSAFGLNRPVLAEDAQRFTALPFPFKKNCNDNICQDDLSITPFSMSLDCLVVG 780
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPFLDLSYKYSTLQORSORSWRLACASASTEV 840
Db 781 GPREFNVTVVRNDGEDSYRTQVTFPFLDLSYKYSTLQORSORSWRLACASASTEV 840
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGKLLKANVTSENNMPTNKTEF 900
Db 841 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGKLLKANVTSENNMPTNKTEF 900
Qy 901 QLELPVKYAVYVWVTSKYNLFTASENTSRVMOHQVSNLQORSIPISLVLVPV 960
Db 901 QLELPVKYAVYVWVTSKYNLFTASENTSRVMOHQVSNLQORSIPISLVLVPV 960
Qy 961 RLNQTVIWDREPQVTFSEKLSSTCKTKERLPSHSDFLAELKAPVYVNCIAVCQRIQCDIP 1020
Db 961 RLNQTVIWDREPQVTFSEKLSSTCKTKERLPSHSDFLAELKAPVYVNCIAVCQRIQCDIP 1020
Qy 1021 PFGIOEEFNATLKNLSFDWIKTSHNHLIIVSTAEILFNDSPVTLPGQAFVRSQTET 1080
Db 1021 PFGIOEEFNATLKNLSFDWIKTSHNHLIIVSTAEILFNDSPVTLPGQAFVRSQTET 1080
Qy 1081 KVEPFEVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGABEQ 1137
Db 1081 KVEPFEVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGABEQ 1137

138 LFVSRQECPRQODIVFLIDSGSSISSENFATMMFVRAVISQORPSTQSLMQFSNKF 197
181 RHFTFKPQNNPNSLVKPTQTLGRTHATANGVKVIRELLNITNGARKNAFKILIVI 240
198 QTHFTFEFRSTNSPLSLASVHQLQGTFTYATATQNVVHRLPHASYGARDAIKILIVI 257
241 TDGKFGDPLGYEDVPEADREGVRYVIGVDAPRSEKSRQELNITASKPRDHVQVN 300
258 TDGKKGSDLDYKVPWADAAGIIRYAIYVGLAFQNSWKNELNDIASQSEHIFKVE 317
301 NFEALKTQNLQREKIPALEGTOTGSSSPHEMSQEGFSAATNSGPLLSTVGSYDAG 360
318 DFDALKQIQNLQREKIPALEGTOTGSSSPHEMSQEGFSAATNSGPLLSTVGSYDAG 377
361 GYFLYTSKESKSTFNNMTVDSMDADYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
378 GAFLYPPNMSPTFINNSQENVMDRSDYLGSTELALWKQSLVGLGAPRYOHIGLVAMFR 437
421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPRYOHIGLVAMFR 480
438 QVSRQRMKAEBVIGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPRYOHIGLVAMFR 497
481 PRGQARWCCDAVLYGEGQHPWGRFGAALTVDLVGVDNGKLTVDVIGAPSEENRGAYLF 540
498 PRGWR-RWVCDVLYGEGQHPWGRFGAALTVDLVGVDNGKLTVDVIGAPSEENRGAYLF 556
541 HGTSGSGISPSHSORIASKSLSPRIQYQSGSLSGGQDLTMDGLVDLVGAGHVLILRSQ 600
557 HGVLGSPISPSHSORIASKSLSPRIQYQSGSLSGGQDLTMDGLVDLVGAGHVLILRSQ 616
601 PVLRYKAIEMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTDRRLREGQIQSVVT 660
617 PVLWVGVMQFIPAEIPRSAFECEQVVSQTLVQSNICLVYIDKRSKNTLGSRLQSSVT 676
661 YDLALDGRPHSRVAFNETKNSRQTVLGLTCTETLKLQLPNCIEDVPSIVLRLNF 720
677 LDALAPGRLSPRAIFQETKNSRVRVGLKAKHCFNLLPSCEVDSVPIILRLNF 736
721 SVLGTPLSAFGLNRPLAEDAQRLFTALFPPEKNCNDNICODDLSITPSFMSLDCLVWG 780
737 TLVKGFLAFLNRLPLAEDAQRYFTASLPPEKNCNDNICODDLSITPSFMSLDCLVWG 796
781 GPRENVTVVRNDEGDSYRQVTPFPPLDLSYKSTVLQNRQSRQWRACASSTEV 840
797 SNLELNARVWVNDGDSYGTITTFSPAGLSYRYVAEGQKQGLRSLHLC--CSAPVG 854
841 SGALKSTSCSINHPIFPNSBVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 900
855 SGTWSTSCSINHPIFRGAQITFLATFDVSPKAVGLDRULLIANVSENNIPRTSKTIF 914
901 QLELPVKYAVVWVTSKSTYKYNLPTAS-BNTRVMOHQVQVNSNLGQSLPISLFLVP 959
915 QLELPVKYAVVWVTSKSTYKYNLPTAS-BNTRVMOHQVQVNSNLGQSLPISLFLVP 974
960 VRLNQTVIWRPQVTFSENLSSTCTKRLPSPHSDPLAELRKAPVNVNCIAVQRIQCDI 1019
975 VELNQAVMDVEVSHPCNPSLRCSSEKIPASDPLAHLQKVPVLDLCSLACGLRFECDV 1034
1020 PFFGLOEFNATLKNLSFDWYKTSNELLIVSTABILFNDVSTFLPCCQGAFFVRSQTE 1079
1035 PSFSVQEELOFTLKNLSFGVVRQIQKQVSVSWABEIIFTSVYQSLPQCEAFMBAQTI 1094
1080 TKVPPFVPEVPELPIVGVSSVGLLALLALITAALYKLGFFKQVQKDWMSH 1128
1095 TVLEKYKVENPIPIVGVSSIGULLLALLITAVLYKVGFFKQYKEMWEE 1143

RESULT 13

US-09-891-943-4

; Sequence 4, Application US/09891943

; Publication No. US2003007279A1

; GENERAL INFORMATION:

; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US2003007278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-15
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-891-943-4

Query Match 58.7%; Score 3450; DB 10; Length 1163;
Best Local Similarity 61.0%; Pred. No. 7.4e-308;
Matches 589; Conservative 144; Mismatches 290; Indels 6; Gaps 4;
QY 1 FNLDTENAMTFOENARGFGQSVWOLQSGSVVVGAPQEIIVANQSGSLYQCOYSTGSCBPI 60
DB 20 FNLDTTEELTAFRVDSAGFGSDVQVYANSVWVVGAPQKIIAANOIGLYQCGYSTGACBPI 79
QY 61 RLQVPEAVNNSLGLSLAATSPQLLACGPTVHTQTCSENTYVKGCLFPLFQSNLRQOQPK 120
DB 80 GLQVPEAVNNSLGLSLASTTSPQLLACGPTVHCEGRNMYLTLGCLFLGPT--QLTOR 137
QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIHIDFRMKEFVSTVMEQKKSKTLFSLMOYSEEP 180
DB 138 LFVSRQECPRQODIVFLIDSGSSISSENFATMMFVRAVISQORPSTQSLMQFSNKF 197
QY 181 RHFTFKPQNNPNSLVKPTQTLGRTHATANGVKVIRELLNITNGARKNAFKILIVI 240
DB 198 QTHFTFEFRSTNSPLSLASVHQLQGTFTYATATQNVVHRLPHASYGARDAIKILIVI 257
QY 241 TDGKFGDPLGYEDVPEADREGVRYVIGVDAPRSEKSRQELNITASKPRDHVQVN 300
DB 258 TDGKKGSDLDYKVPWADAAGIIRYAIYVGLAFQNSWKNELNDIASQSEHIFKVE 317
QY 301 NFEALKTQNLQREKIPALEGTOTGSSSPHEMSQEGFSAATNSGPLLSTVGSYDAG 360
DB 318 DFDALKQIQNLQREKIPALEGTOTGSSSPHEMSQEGFSAATNSGPLLSTVGSYDAG 377
QY 361 GYFLYTSKESKSTFNNMTVDSMDADYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
DB 378 GAFLYPPNMSPTFINNSQENVMDRSDYLGSTELALWKQSLVGLGAPRYOHIGLVAMFR 437
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPRYOHIGLVAMFR 480
DB 438 QVSRQRMKAEBVIGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPRYOHIGLVAMFR 497
QY 481 PRGQARWCCDAVLYGEGQHPWGRFGAALTVDLVGVDNGKLTVDVIGAPSEENRGAYLF 540
DB 498 PRGWR-RWVCDVLYGEGQHPWGRFGAALTVDLVGVDNGKLTVDVIGAPSEENRGAYLF 556
QY 541 HGTSGSGISPSHSORIASKSLSPRIQYQSGSLSGGQDLTMDGLVDLVGAGHVLILRSQ 600
DB 557 HGVLGSPISPSHSORIASKSLSPRIQYQSGSLSGGQDLTMDGLVDLVGAGHVLILRSQ 616
QY 601 PVLRYKAIEMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTDRRLREGQIQSVVT 660
DB 617 PVLWVGVMQFIPAEIPRSAFECEQVVSQTLVQSNICLVYIDKRSKNTLGSRLQSSVT 676
QY 661 YDLALDGRPHSRVAFNETKNSRQTVLGLTCTETLKLQLPNCIEDVPSIVLRLNF 720
DB 677 LDALAPGRLSPRAIFQETKNSRVRVGLKAKHCFNLLPSCEVDSVPIILRLNF 736

QY	721	SLVGTPLSAGNLRPVLAEDAQRFLTALPFEPFKNCGNDNICDDLSITPFSMELDCLVVG	760
Db	737	TLVGPILLAPRNLRPMLAALAQRVFTASLPFEKNCADHICQDNLGISFSPGLKSLVVG	796
QY	781	GPREFNVTVVRDGDSDSYRTQVTFPPFLDLSVRKYSTLQNSQRSWRWLACASASSTEV	840
Db	797	SNLELNAEVMVNDGSDSYGTTTFHPAGLSVRYVAEGQKQGLSLHUTC--CSAPVG	854
QY	841	SGALKTSCSINHPITPENSESEVFNTITFDVDSKASLGNKLLKANVTSENMMPTNKTEF	900
Db	855	SQGTWTSICRINHILFRGGAQITFLATFDVSPRAVGLDRLLLIANYSSENNIPTSKTIF	914
QY	901	QLSLPVKYAVMVVTVSHGVSTKYLNFITAS-ENTSRVMQHOYQVSNLQORSPLSLVPLVP	959
Db	915	QLSLPVKYAVYIVVSSEHQFTKYLNFSESEKSHVAMHRYQVNNLQORLDLPVSNPMPV	974
QY	960	VLRLNQTVMIDRPQVTFSENLSSTCHTKERLPSDSFLAELRKAPVNVNCSIAVCQRICQDI	1019
Db	975	VELNQEAVMVNDVEVSPHNPQNSSEKIPASDFLAHTQKNPVLDCSIAQCLRFPCDV	1034
QY	1020	PFGIOBEFNATUKGNLSFOWYIKTSHNHLLIYSTABILFNDVSFTLLPQGGAPVRSQTE	1079
Db	1035	PSFSVQBELDFTLKENLSPQWVRQILQKQVSVSVABEIIFDTSYISOLPQGEAFMRAQTI	1094
QY	1080	TKYPEPEVPNPLPLIVGSSVGGELLLALITAAALYKLGFEKROYKDMNSE	1128
Db	1095	TVLEKYKVRNPIPLIVGSSIGLLLLALITAVLYKVGFFRQYKEMMEZ	1143
RESULT 14			
US-09-350-259-2			
; Sequence 2, Application US/09350259			
; Patent No. US20020062008A1			
; GENERAL INFORMATION:			
; APPLICANT: Gallatin, Michael W.			
; APPLICANT: Van der Vieren, Monica			
; TITLE OF INVENTION: No. US20020062008A1el Human 2			
; FILE REFERENCE: 27866/35004			
; CURRENT APPLICATION NUMBER: US/09/350,259			
; CURRENT FILING DATE: 1993-07-08			
; EARLIER APPLICATION NUMBER: 09/193,043			
; EARLIER FILING DATE: 1998-11-16			
; EARLIER APPLICATION NUMBER: 08/173,497			
; EARLIER FILING DATE: 1993-12-23			
; EARLIER APPLICATION NUMBER: 08/286,889			
; EARLIER FILING DATE: 1994-08-05			
; EARLIER APPLICATION NUMBER: 08/362,652			
; EARLIER FILING DATE: 1994-12-21			
; EARLIER APPLICATION NUMBER: 08/943,363			
; EARLIER FILING DATE: 1997-10-03			
; NUMBER OF SEQ ID NOS: 114			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 2			
; LENGTH: 1161			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-350-259-2			
Query Match 58.2%; Score 3417; DB 9; Length 1161;			
Best Local Similarity 59.8%; Pred. No. 8, 1e-305;			
Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;			
QY	1	FNLDTEANMTQENARGFGQSVVLQGSRRVVGPQEIIVANQRGSLVQCTYSTGSCPEI	60
Db	17	FNLDVEEPTIFQEDAGGFGQSVVFGGSRVVGAPLEVVAAVAAQTERLDYCAAAATGWCQPI	76
QY	61	PLQYVEAVNMSLGLSLAATYSPOLLACGPTVHQTCSNTYVKGCLCFLFGSNLRQOPK	120
Db	77	PLHIRPEAVNMSLGLTAASTNGSKLLACGPTVHRVCGENSYKSGSCULLLSRW-EIQT	135
QY	121	FPEALRGCTQBDSDIAFLIDGSGSIIPDFRMRKEFVSTVMEQLKSKTFLPSLMQYSEEF	180

136	VPDAPTECHQEMDWIVFLIDSGSSIDQNDPQMKQFVQAVMCQFEGTDTLTFALMQVSNLL	195
181	RIHFTPKFQNNPNRSLVKPITQOLLGRTHPTATGKRVITRELLNITNGARKQAFKILVI	240
196	KIHFTTQFRTSPSQOQLDPIVOLKGLTFTATGLTVVTQFHEKNGARKSAKKILVI	255
241	TDCEKFGDPLGVEDVITPEADREGVTRYVITGVGDAPRSSEKSRQBLNITIASKPRDRHVQVN	300
256	TDQCKYKDPLEYSDVITPOAEKAGIIRYALGVGHATQGPRTARQELNITISSAPQDHVFKVD	315
301	NPEALKTIQNLREKIPAJEGTQGTSSSFHEHMQSGEFSAITNSGPISTTVSGVSDWAG	360
316	NFAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQBGFTALTMDGLFLGAVGFSWSG	375
361	GVPLYTSKEKSTFINNTRVDSDMNDAYLGAAGIILNRNVQSILVCAPRYQHIGLVAMFR	420
376	GATLYPNNMPTFINMSQENVDSDSYLSYSELALWKGVQNLVLGAPRYQHTGKAVIF	435
421	QNTGMWESNANVKTQIGAYFGASLCSVDVDNSGSDTLVLIGAPHYETROGQVSVCLP	480
436	QVSRQWRKKAETVGTQIGSYFGASLCSVDVDSDGSDTLVLIGAPHYETROGQVSVCLP	495
481	PRGORARWQDAVLYGEOQCPNCRFGAALTVLGDVNGDKLTDVAICAPGEENRCAVILP	540
496	PRGORVQWQDAVLRQEGQHPWRFGAALTVLGDVNEDKLIDVAICAPGEENRCAVILP	555
541	HGTSGSGISPSHQRAGSKLSPRLQYFQOSLGGQDLTMDGLVDLTVGAQGHVILLRSQ	600
556	HGASESGISPSHSORIASQSLSPRLQYFQOALSGGQDLTQDGLMDLAVARGOVLILRSL	615
601	PVLRYKAIEMFNPREVARNVPEQNDQVVKGEAGVRYCLHVQKSTRDRILREGQIOSVVT	660
616	PVLKVGVAKEFSPVYVAKAVYRCWEXPSALEAGDATVCLTIOKSSLDQL--GDIOSSVR	673
661	YDLALDSRPHSRVAFNETKNSRTRRTQVLEGLTQTCETLKLQLPNCIBDPSVPIVLRNP	720
674	FDLALDPGLTSLRAIPNETKNPILTRRKTLGLGIHCETLLKLPLDCVBDVWVSPILHLNF	733
721	SLVCTPLUSAFGNLPUAEDAQRLFTALPPEKXNGNDNICODDLSTPSEKSLDCLVWG	780
734	SLVREPIPSQNLRPVLAVGSDQLFTASLPEFKNCQDGLCEBGLGVTLSFSGLOTLTVG	793
781	GPREFNVTVVRNDGDSYRTQVTFPPFLDLSYRKYSTTLQNBRSQSRWRLACSSASTEV	840
794	SSLELVIVTVWNRAGEDSYCTVWSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED	852
841	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKLLKANTVISENNMERTKTFE	900
853	EG-LRSSRCSVNHPIFHEGNGTFITFDVDSYKATLGDRLMMKASSEKANKASSKATF	911
901	QLELPVKYAVVMVYTHSGVSTKYLNP--TASENTSRVMQHOYQVSNLQSRSLPISLVFLVP	959
912	QLELPVKYAVYTHLSRQEBSTKYFNFATSDKKWKEAEHRYRVNNLSQRDLAISINFVWP	971
960	VRNLQTVIMDRPQVTFSENISSCTCHKTERLPSHSDFLAEIRKAPVNVCSLAVCORIQCDI	1019
972	VLLNGVAVMDVMEAPSQSL--PCVSRPKPPQHSDFLTOISRSPMLDCSTADCLQPRCDV	1029
1020	PFGIOREAFNATLKGNIISPDWYIKTSHNHLIIVSTAEILENDVSVFTLLPQGAFFVRSQTE	1079
1030	PSFSVQBELDFTLKGNIISFGWRETLQCKVLVSVASIEITDTSVYSQLPQEAFFMRAQME	1089
1080	TKVPEFVNPPLPIVSGSVGLLILLALITAALYKLGFFKQKYDKMMSE	1128
1090	MVLSEDEVNAPLPIINGSSVGLLILLALITATLYKLGFFKRYKEMLED	1138

RESULT 15
US-09-891-943-2 ; Sequence 2, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.

APPLICANT: Van der Vieren, Monica	
TITLE OF INVENTION: No. US20030077278A1e1 Human 2	
FILE REFERENCE: 27866/35004	
CURRENT APPLICATION NUMBER: US/09/891,943	
CURRENT FILING DATE: 2001-06-26	
PRIOR APPLICATION NUMBER: 09/193,043	
PRIOR FILING DATE: 1998-11-16	
PRIOR APPLICATION NUMBER: 08/286,889	
PRIOR FILING DATE: 1994-08-05	
PRIOR APPLICATION NUMBER: 08/362,652	
PRIOR FILING DATE: 1994-12-21	
PRIOR APPLICATION NUMBER: 08/943,363	
PRIOR FILING DATE: 1997-10-03	
NUMBER OF SEQ ID NOS: 114	
SOFTWARE: PatentIn Ver. 2.0	
SEQ ID NO 2	
LENGTH: 1161	
TYPE: PRT	
ORGANISM: Homo sapiens	
US-09-891-943-2	
Query Match	
Best Local Similarity 59.8%; Pred. No. 8.1e-305;	
Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;	
Qy	1 FNLDTENAMTFOENARCFQSVVLOGSRVVGAPQEIYAANQSGSLYQCDYSTGSCPEI 60
Db	17 FNLDVEEPTTFQEDAGFGQSVQVQGGSLRVVGAPELVAAANTQGRVYDCAATGHCQPI 76
Qy	61 RLQVPVAVNMSGLSLAATTSPPQLACGFTVHQTCSENTRYKGLCFLFGSNLRQPOK 120
Db	77 PLHIRPEAVNMSGLTAASTNGSRLLACGFTLHRVCGENSYSKSGCLLILGSRW-BIIQT 135
Qy	121 FPEALRCQPEDSDIAFLIDGSGIIPHDPRMKEFYSTWMEQLKSKTLFSLMOYSEEP 180
Db	136 VPDATPCPHQEMDIVFLIDGSGIDQNDENQMKGFVQAVNGQFEGFTLFLALMQSYNLL 195
Qy	181 RIHFTFEFQNPNSRLVXPIITQLLGRTHATGVRIRELLNITNGAKRNAFKILIVI 240
Db	196 KIHFTFTQFTSPSQSVDPVIVQLKGLTFTATGILTVVTLQFHFKNGARKSAKKILIVI 255
Qy	241 TDGKFGDPLGYEDVPEADREGVIRVVGVDAPFRSEKSRQELNITIASKPRDHVQVN 300
Db	256 TDGQKYDPLEYSDVPOAEKAGIIRYAIGVGHAFQGTARQELNITISSAPPQDHVEKVD 315
Qy	301 NFEALKTIQNLREKI PAIEGTQTGSSSFEHENSQGFSAATISNGPILLSVGSYDWAG 360
Db	316 NFAALGSIQXQLQEKIYAVEGTQGRASSPQFEMSQGFSTALTMQGLFLGAVGSPFSG 375
Qy	361 GVFLYTSKEKSTFINMTRVDSMDKDAVIGYAAAIIILNRVQSULVIGAPRYOHIGLVAMFR 420
Db	376 GAFLYPPNMSPTFINMSQENVMDRDSYLGYSTELALAKGVQNLVIGAPRYOHTGKAVIFT 435
Qy	421 QNTGWESNANVKTQIGAVFGASICSVDVDSNGSTDLVILGAPHYVQTRGGQVSVCP 480
Db	436 QVSRQWRKKAFTVGTQIGSYFGASLCSVDVDSGSDTDLILGAPHYVQTRGGQVSVCP 495
Qy	481 PRGQARWQCDAVLYGECQFWRFGAALTVLGVNGDKLTVAIGAPEGEDNRGAVYLF 540
Db	496 PRGQVQVQCCDAVLRGEGHFWGRFGAALTVLGDVNEBKLDVAIGAPEGQENRGAVYLF 555
Qy	541 HGTSGGSI SPHSQRIASKLSPRIQYFGQSLSGGDLTMQGLVDLTVGAGHVLILRSQ 600
Db	556 HGASESGI SPHSQRIASSQLSPRIQYFGQALSQGGDLTQDGLMDLAVGARGQVILLRSL 615
Qy	601 PVLRYKALMEFNPREVARNVPECNDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVWT 660
Db	616 PVLKVGVARMSPEVAAYKCEWEEKPSALENGDATVCLTIQKSSLDL--GDIQSSVR 673
Qy	661 YDLALDSGRPHSRVFNETHKSTRTQTVIGLGTQTCETLKLQLCNIEDPVPVILRLNF 720
Db	674 PDLALDPGLRTSRAIFNETKNPTLTKRTKTLGLIGHCTLKLILLPCDVEDVVPVILRLNF 733

Search completed: June 7, 2004, 17:38:55
Job time : 42.8905 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:07:20 ; Search time 18.3484 Seconds
(without alignments)
3199.127 Million cell updates/sec

Title: US-09-902-481B-6
Perfect score: 5875
Sequence: 1 FNLDTENAMTFOENARGFGQ.....FKROYKDMSEGGPGGAEPO 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.psp:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.psp:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.psp:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.psp:*
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.psp:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.psp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5868	99.9	1153	1 US-08-173-497-3	Sequence 3, Appli
2	5868	99.9	1153	1 US-08-286-889-3	Sequence 3, Appli
3	5868	99.9	1153	1 US-08-485-618-3	Sequence 3, Appli
4	5868	99.9	1153	1 US-08-362-652-3	Sequence 3, Appli
5	5868	99.9	1153	2 US-08-605-672-3	Sequence 3, Appli
6	5868	99.9	1153	2 US-08-482-293A-3	Sequence 3, Appli
7	5868	99.9	1153	2 US-08-943-363-3	Sequence 3, Appli
8	5868	99.9	1153	3 US-09-193-043-3	Sequence 3, Appli
9	5868	99.9	1153	4 US-09-688-307A-3	Sequence 3, Appli
10	5868	99.9	1153	4 US-09-350-259-3	Sequence 3, Appli
11	5837.5	99.4	1152	2 US-08-476-062A-43	Sequence 43, Appl
12	5837.5	99.4	1152	5 PCT-US96-01314-43	Patent No. 5424399
13	5837.5	99.4	1152	6 US-08-476-062A-44	Sequence 44, Appl
14	3473	59.1	1163	2 PCT-US96-01314-44	Sequence 44, Appl
15	3473	59.1	1163	1 US-08-173-497-4	Sequence 4, Appli
16	3450	58.7	1163	1 US-08-286-889-4	Sequence 4, Appli
17	3450	58.7	1163	1 US-08-485-618-4	Sequence 4, Appli
18	3450	58.7	1163	1 US-08-362-652-4	Sequence 4, Appli
19	3450	58.7	1163	1 US-08-605-672-4	Sequence 4, Appli
20	3450	58.7	1163	2 US-08-482-293A-4	Sequence 4, Appli
21	3450	58.7	1163	2 US-08-943-363-4	Sequence 4, Appli
22	3450	58.7	1163	3 US-09-193-043-4	Sequence 4, Appli
23	3450	58.7	1163	4 US-09-688-307A-4	Sequence 4, Appli
24	3450	58.7	1163	4 US-09-350-259-4	Sequence 4, Appli
25	3450	58.7	1163	4 US-08-173-497-2	Sequence 2, Appli
26	3417	58.2	1161	1 US-08-286-889-2	Sequence 2, Appli
27	3417	58.2	1161	1 US-08-485-618-2	Sequence 2, Appli

28	3417	58.2	1161	1 US-08-485-618-2	Sequence 2, Appli
29	3417	58.2	1161	1 US-08-362-652-2	Sequence 2, Appli
30	3417	58.2	1161	2 US-08-605-672-2	Sequence 2, Appli
31	3417	58.2	1161	2 US-08-482-293A-2	Sequence 2, Appli
32	3417	58.2	1161	2 US-08-943-363-2	Sequence 2, Appli
33	3417	58.2	1161	3 US-09-193-043-2	Sequence 2, Appli
34	3417	58.2	1161	4 US-09-688-307A-2	Sequence 2, Appli
35	3417	58.2	1161	4 US-09-350-259-2	Sequence 2, Appli
36	3401.5	57.9	1161	1 US-08-485-618-99	Sequence 99, Appl
37	3401.5	57.9	1161	2 US-08-605-672-99	Sequence 99, Appl
38	3401.5	57.9	1161	2 US-08-482-293A-99	Sequence 99, Appl
39	3401.5	57.9	1161	2 US-08-943-363-99	Sequence 99, Appl
40	3401.5	57.9	1161	3 US-09-193-043-99	Sequence 99, Appl
41	3401.5	57.9	1161	4 US-09-688-307A-99	Sequence 99, Appl
42	3401.5	57.9	1161	4 US-09-350-259-99	Sequence 99, Appl
43	3239.5	55.1	1161	3 US-09-193-043-55	Sequence 55, Appl
44	3239.5	55.1	1161	4 US-09-688-307A-55	Sequence 55, Appl
45	3239.5	55.1	1161	4 US-09-350-259-55	Sequence 55, Appl

ALIGNMENTS

RESULT 1
US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-3

Query Match 99.9%; Score 5868; DB 1; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 1; Mismatches 0; Gaps 0;

OY 1 FNLDTENAMTFOENARGFGQSVVVGAEPOEIVAAANQGSLSYCCDYSTGSCSEPI 60

DB 17 FNLDTENAMTFOENARGFGQSVVVGAEPOEIVAAANQGSLSYCCDYSTGSCSEPI 76

QY 61 RLQVPEAVNMSGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFLPGSNLRQPOK 120
DB 77 RLQVPEAVNMSGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFLPGSNLRQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEOLKXSKTLFSLMOYSEEP 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEOLKXSKTLFSLMOYSEEP 196
QY 181 RIHFTFKFQNNPNRSLVKPIITOLLGRTHATGVRKVIKRELLNITNGARKNAFKILVI 240
DB 197 RIHFTFKFQNNPNRSLVKPIITOLLGRTHATGVRKVIKRELLNITNGARKNAFKILVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQBLNTIASKPPRDHVFQVN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQBLNTIASKPPRDHVFQVN 316
QY 301 NFEALKTIONQLREKI FAIEGTQTGSSSEHEMSQEGFAAITSNGPLISTVGSYDAG 360
DB 317 NFEALKTIONQLREKI FAIEGTQTGSSSEHEMSQEGFAAITSNGPLISTVGSYDAG 376
QY 361 GVFLYTSKESKSTFNMTVRSDMDNDAYLGAAAIILNRRVQSLVIGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKESKSTFNMTVRSDMDNDAYLGAAAIILNRRVQSLVIGAPRYOHIGLVAMFR 436
QY 421 QNTGWESNANVKGTQICAYFGASLCSVDVDSNGSTDLVIGAPHYTEOTRGQVSCPL 480
DB 437 QNTGWESNANVKGTQICAYFGASLCSVDVDSNGSTDLVIGAPHYTEOTRGQVSCPL 496
QY 481 PRGRARWQCDVLYGCGQGWGFAGALTVLGDVNGDKLTDVAIGAPGEEDNEGAVYLP 540
DB 497 PRGRARWQCDVLYGCGQGWGFAGALTVLGDVNGDKLTDVAIGAPGEEDNEGAVYLP 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTWDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTWDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLAVKALMEPNPREVARNVPCNDQVVKGEACEVAVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLAVKALMEPNPREVARNVPCNDQVVKGEACEVAVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDGRPHSRVAFNFKNSTRQRTQVGLGTQTCBTLKLQLPNCIEDPVPVILRLNF 720
DB 677 YDLALDGRPHSRVAFNFKNSTRQRTQVGLGTQTCBTLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLFTALFPPEKNCNDNICQDDLSTTFSPMSLDCLVVG 780
DB 737 SLVGTPLSAFGLNRPVLAEDAQRLFTALFPPEKNCNDNICQDDLSTTFSPMSLDCLVVG 796
QY 781 GPRFNVTVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASTEV 840
DB 797 GPRFNVTVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVTNIPFDVDSKASLGKLLKANVTSENNMPTNKTEP 900
DB 857 SGALKSTSCSINHPIPPENSEVTNIPFDVDSKASLGKLLKANVTSENNMPTNKTEP 916
QY 901 QLELPVKYAVVWVTSKYNLFTASNTSRVMOHQYQVSNLQORSIPISLVFLVPV 960
DB 917 QLELPVKYAVVWVTSKYNLFTASNTSRVMOHQYQVSNLQORSIPISLVFLVPV 976
QY 961 RLNQTIVWRDQVTPSENLSTCTTKERLPKSHSDFLAEKRAKAPVNCISVQRIQCDIP 1020
DB 977 RLNQTIVWRDQVTPSENLSTCTTKERLPKSHSDFLAEKRAKAPVNCISVQRIQCDIP 1036
QY 1021 PFGIOEEFNATLKGNSFDWIKTSHNHLIVSTAEILFNDSVFTLLPGQAPVRSQET 1080
DB 1037 PFGIOEEFNATLKGNSFDWIKTSHNHLIVSTAEILFNDSVFTLLPGQAPVRSQET 1096
QY 1081 KYPEPEVFNPLPIVGVSSVGLLLAL:TAALYKLGFFKRYKQKDMWSEGGPPGAEPQ 1137
DB 1097 KYPEPEVFNPLPIVGVSSVGLLLAL:TAALYKLGFFKRYKQKDMWSEGGPPGAEPQ 1153

RESULT 2
US-08-286-889-3
Sequence 3, Application US/08286889
Patent No. 5470953
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Mich
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murtay & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,889
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: P38,659
REFERENCE/DOCKET NUMBER: 27866/32168
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-286-889-3

Query Match 99.9%; Score 5868; DB 1; Length 1153;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFQENARFGQSVVQLQGSRVVVGAPQEIIVAAQORGSLYQCDYSTGCEPI 60
DB 17 FNLDTENAMTFQENARFGQSVVQLQGSRVVVGAPQEIIVAAQORGSLYQCDYSTGCEPI 76
QY 61 RLQVPEAVNMSGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFLPGSNLRQPOK 120
DB 77 RLQVPEAVNMSGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFLPGSNLRQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEOLKXSKTLFSLMOYSEEP 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEOLKXSKTLFSLMOYSEEP 196
QY 181 RIHFTFKFQNNPNRSLVKPIITOLLGRTHATGVRKVIKRELLNITNGARKNAFKILVI 240
DB 197 RIHFTFKFQNNPNRSLVKPIITOLLGRTHATGVRKVIKRELLNITNGARKNAFKILVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQBLNTIASKPPRDHVFQVN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQBLNTIASKPPRDHVFQVN 316
QY 301 NFEALKTIONQLREKI FAIEGTQTGSSSEHEMSQEGFAAITSNGPLISTVGSYDAG 360
DB 317 NFEALKTIONQLREKI FAIEGTQTGSSSEHEMSQEGFAAITSNGPLISTVGSYDAG 376

361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRRVQSLVLAGPRYOHIGLVAMFR 420
Db GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRRVQSLVLAGPRYOHIGLVAMFR 436
377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRRVQSLVLAGPRYOHIGLVAMFR 480
Db GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRRVQSLVLAGPRYOHIGLVAMFR 496
421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQGVSCPL 540
Db QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQGVSCPL 556
437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQGVSCPL 600
Db QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQGVSCPL 616
481 PRGORARWQCDVLYGEGQGWGFGAALTVLGDNVGDCLTDVAIGAPGEDNRGAVYLF 660
Db PRGORARWQCDVLYGEGQGWGFGAALTVLGDNVGDCLTDVAIGAPGEDNRGAVYLF 676
497 PRGORARWQCDVLYGEGQGWGFGAALTVLGDNVGDCLTDVAIGAPGEDNRGAVYLF 720
Db PRGORARWQCDVLYGEGQGWGFGAALTVLGDNVGDCLTDVAIGAPGEDNRGAVYLF 736
541 HGTSGSGISPHSQRISAGSKLSPRLQVFGQSLSGGQDLTMDGLVDTLVGAQGHVLLRSQ 780
Db HGTSGSGISPHSQRISAGSKLSPRLQVFGQSLSGGQDLTMDGLVDTLVGAQGHVLLRSQ 796
557 HGTSGSGISPHSQRISAGSKLSPRLQVFGQSLSGGQDLTMDGLVDTLVGAQGHVLLRSQ 840
Db HGTSGSGISPHSQRISAGSKLSPRLQVFGQSLSGGQDLTMDGLVDTLVGAQGHVLLRSQ 856
601 FVLRAVKAIMEFNPREVARNVPCNDQVVKGEAGEVRVCLRVOKSTRDLRREGQIOSVVT 900
Db FVLRAVKAIMEFNPREVARNVPCNDQVVKGEAGEVRVCLRVOKSTRDLRREGQIOSVVT 916
661 YDLALDSGRPSRAVNETKSTRQVGLTQTCETLKLQLPNCIEDPVSPIVLRNP 960
Db YDLALDSGRPSRAVNETKSTRQVGLTQTCETLKLQLPNCIEDPVSPIVLRNP 976
677 YDLALDSGRPSRAVNETKSTRQVGLTQTCETLKLQLPNCIEDPVSPIVLRNP 1020
Db YDLALDSGRPSRAVNETKSTRQVGLTQTCETLKLQLPNCIEDPVSPIVLRNP 1036
721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKXGNDNICODDLSITFSFMSLDCLVG 1080
Db SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKXGNDNICODDLSITFSFMSLDCLVG 1096
737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKXGNDNICODDLSITFSFMSLDCLVG 1137
Db SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKXGNDNICODDLSITFSFMSLDCLVG 1153
781 GPREFNVTVRNDGDSYQTQVTFPPPLDLSYKXVSTLQNSORSWRACLSASTEV 840
Db GPREFNVTVRNDGDSYQTQVTFPPPLDLSYKXVSTLQNSORSWRACLSASTEV 856
797 GPREFNVTVRNDGDSYQTQVTFPPPLDLSYKXVSTLQNSORSWRACLSASTEV 900
Db GPREFNVTVRNDGDSYQTQVTFPPPLDLSYKXVSTLQNSORSWRACLSASTEV 916
841 SGALKSTSCSINHPIFPENSEVFNIFDVSXASLGNKLLKXANTSENNPRNTKTEF 960
Db SGALKSTSCSINHPIFPENSEVFNIFDVSXASLGNKLLKXANTSENNPRNTKTEF 976
901 QLELPVYAVVMVTSKGVSTKYLNFTASENTSRVMOHQVQNSLNGORSPLISLVLVFP 1020
Db QLELPVYAVVMVTSKGVSTKYLNFTASENTSRVMOHQVQNSLNGORSPLISLVLVFP 1036
917 QLELPVYAVVMVTSKGVSTKYLNFTASENTSRVMOHQVQNSLNGORSPLISLVLVFP 1080
Db QLELPVYAVVMVTSKGVSTKYLNFTASENTSRVMOHQVQNSLNGORSPLISLVLVFP 1096
961 RLNQTIVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKPVWNCSTAVCORIOCDIP 1137
Db RLNQTIVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKPVWNCSTAVCORIOCDIP 1153
977 RLNQTIVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKPVWNCSTAVCORIOCDIP 1137
Db RLNQTIVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKPVWNCSTAVCORIOCDIP 1153
1021 PFGIOEFNATLKNLSFDWYKTSNHLIIVSTAEILFNDVSTLPLPGQAFVRSQTTET 1080
Db PFGIOEFNATLKNLSFDWYKTSNHLIIVSTAEILFNDVSTLPLPGQAFVRSQTTET 1096
1037 PFGIOEFNATLKNLSFDWYKTSNHLIIVSTAEILFNDVSTLPLPGQAFVRSQTTET 1137
Db PFGIOEFNATLKNLSFDWYKTSNHLIIVSTAEILFNDVSTLPLPGQAFVRSQTTET 1153
1081 KVEPFEVNPPLIVGSSVGLLILALITAAALYKLGFFKQYKDMWSEGGPPGAEPO 1137
Db KVEPFEVNPPLIVGSSVGLLILALITAAALYKLGFFKQYKDMWSEGGPPGAEPO 1153
1097 KVEPFEVNPPLIVGSSVGLLILALITAAALYKLGFFKQYKDMWSEGGPPGAEPO 1137
Db KVEPFEVNPPLIVGSSVGLLILALITAAALYKLGFFKQYKDMWSEGGPPGAEPO 1153

RESULT 3

US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08485,618
APPLICATION NUMBER: US/08485,618
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32797
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: Single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-618-3

Query Match 99.9%; Score 5868; DB 1; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLDENAMTPQENARFGQSVVQLQGRVVGAPQIBVAANRGSLYQCDYSTGSCPEI 60
Db 17 FNLDENAMTPQENARFGQSVVQLQGRVVGAPQIBVAANRGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVAVNMSLGLSLAATTPSPOLLACGTPVHOTCENTYVKGCLFGLGSLNRQOPQ 120
Db 77 RLQVPVAVNMSLGLSLAATTPSPOLLACGTPVHOTCENTYVKGCLFGLGSLNRQOPQ 136
QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDPRMKBFVSTVMEQLKSKTLFSLMQYSEEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDPRMKBFVSTVMEQLKSKTLFSLMQYSEEF 196
QY 181 RIHFTFEFQNNPNSILVKPIITOLLGRTHATGVRKVIRESLNTITNGASKNAPKILVI 240
Db 197 RIHFTFEFQNNPNSILVKPIITOLLGRTHATGVRKVIRESLNTITNGASKNAPKILVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDADFSEKSRQBELNTITASKPPRDHVFQVN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDADFSEKSRQBELNTITASKPPRDHVFQVN 316
QY 301 NFRAKTIQNLREKIPIAEGTGTGSSSEFHEMSQEGFSAATISNGPLSTVGSYDNG 360
Db 317 NFRAKTIQNLREKIPIAEGTGTGSSSEFHEMSQEGFSAATISNGPLSTVGSYDNG 376
QY 361 GVPLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRRVQSLVLAGPRYOHIGLVAMFR 420
Db 377 GVPLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRRVQSLVLAGPRYOHIGLVAMFR 436
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQGVSCPL 480
Db 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQGVSCPL 496
QY 481 PRGORARWQCDVLYGEGQGWGFGAALTVLGDNVGDCLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGORARWQCDVLYGEGQGWGFGAALTVLGDNVGDCLTDVAIGAPGEDNRGAVYLF 556
QY 541 HGTSGSGISPHSQRISAGSKLSPRLQVFGQSLSGGQDLTMDGLVDTLVGAQGHVLLRSQ 600

557 HGTSGSGISPHSQRAGSKLSPRLQYFGQSLGGQDLTMDGLVLTGCAQGHVLLRSQ 616
601 PVLRVKAIWEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTEDRLREGQIOSVVT 660
617 PVLRVKAIWEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTEDRLREGQIOSVVT 676
661 YDLALDSGRPHSRAVNETKNSRTRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
677 YDLALDSGRPHSRAVNETKNSRTRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
721 SLVGTPLSAPGNLRPVLAEDAQRLLFTALPPFEXKNCNDNICDDLSITFSMSLDCLVVG 780
737 SLVGTPLSAPGNLRPVLAEDAQRLLFTALPPFEXKNCNDNICDDLSITFSMSLDCLVVG 796
781 GPREFNVTVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSORSQSWRLACESASSTEV 840
797 GPREFNVTVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSORSQSWRLACESASSTEV 856
841 SGALKSTSCSINHPIPPENSEVTFTNTPDVSASLGNKLLKANTYSENMPRTNKTFF 900
857 SGALKSTSCSINHPIPPENSEVTFTNTPDVSASLGNKLLKANTYSENMPRTNKTFF 916
901 QLELPVKAYVMVTSKGVSTKYNFTASNTSRVMQHOYOVNLSGORSIPISLVFLVPV 960
917 QLELPVKAYVMVTSKGVSTKYNFTASNTSRVMQHOYOVNLSGORSIPISLVFLVPV 976
961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNCSTAVCQRIQCDIP 1020
977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNCSTAVCQRIQCDIP 1036
1021 FFGIOEPNATLKNLSPDVYKTSNHLIIVSTABILENDSVFTLLPGOGAFVRSQDET 1080
1037 FFGIOEPNATLKNLSPDVYKTSNHLIIVSTABILENDSVFTLLPGOGAFVRSQDET 1096
1081 KVEPFEVNPPLPLTVGSSVGGILLALITAAALKLGFKKQYKDMMSSEGGPPCAEPQ 1137
1097 KVEPFEVNPPLPLTVGSSVGGILLALITAAALKLGFKKQYKDMMSSEGGPPCAEPQ 1153

RESULT 4
US-08-362-652-3
Sequence 3, Application US/08362652
Patent No. 5766850
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vlieten, Monica
TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,652
FILING DATE:
CLASSIFICATION: 435
PRIORITY INFORMATION:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIORITY INFORMATION:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32391
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-362-652-3
Query Match 99.9%; Score 5868; DB 1; length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLDTENAMTPOENARGFGQSVVQLQSSRVVVGAPQELVAANQORSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTPOENARGFGQSVVQLQSSRVVVGAPQELVAANQORSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHTCSENTYVKGCLFGLNLRQPOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHTCSENTYVKGCLFGLNLRQPOK 136
QY 121 PPEARLGCPOEDSDIAPLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTLPFLMOYSEEF 180
DB 137 PPEARLGCPOEDSDIAPLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTLPFLMOYSEEF 196
QY 181 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIRELINLTNGARKNAFKILIVI 240
DB 197 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIRELINLTNGARKNAFKILIVI 256
QY 241 TDCHEKPGDPLGHEDVPEADREGVIRYVIGVGDAPRSEKSRQELNLTASKPRDHVQFN 300
DB 257 TDCHEKPGDPLGHEDVPEADREGVIRYVIGVGDAPRSEKSRQELNLTASKPRDHVQFN 316
QY 301 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPIILSTVGSYDWAG 360
DB 317 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPIILSTVGSYDWAG 376
QY 361 GVPLTYSKEKSTFINNTRVDSMDNDAYLGYAAAIIRNRVQSLVLCAPYQHIGLVAMER 420
DB 377 GVPLTYSKEKSTFINNTRVDSMDNDAYLGYAAAIIRNRVQSLVLCAPYQHIGLVAMER 436
QY 421 QNTGWESNANVKGTOIGAYFCASLCSVDVDSNGSTDLVLIGAPHYVETRGQVSVCP 480
DB 437 QNTGWESNANVKGTOIGAYFCASLCSVDVDSNGSTDLVLIGAPHYVETRGQVSVCP 496
QY 481 PRGQARWQCDVLYGEOQPMGRFGAALTIVLGDVNGDKLTQVAIGAPGEEENRGAVYLF 540
DB 497 PRGQARWQCDVLYGEOQPMGRFGAALTIVLGDVNGDKLTQVAIGAPGEEENRGAVYLF 556
QY 541 HGTSGSGISPHSQRAGSKLSPRLQYFGQSLGGQDLTMDGLVLTGCAQGHVLLRSQ 600
DB 557 HGTSGSGISPHSQRAGSKLSPRLQYFGQSLGGQDLTMDGLVLTGCAQGHVLLRSQ 616
QY 601 PVLRVKAIWEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTEDRLREGQIOSVVT 660
DB 617 PVLRVKAIWEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTEDRLREGQIOSVVT 676
QY 661 YDLALDSGRPHSRAVNETKNSRTRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
DB 677 YDLALDSGRPHSRAVNETKNSRTRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAPGNLRPVLAEDAQRLLFTALPPFEXKNCNDNICDDLSITFSMSLDCLVVG 780
DB 737 SLVGTPLSAPGNLRPVLAEDAQRLLFTALPPFEXKNCNDNICDDLSITFSMSLDCLVVG 796
QY 781 GPREFNVTVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSORSQSWRLACESASSTEV 840

Db 797 GPREFNTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORSQBSWELACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGNKLLKANKVTSNNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGNKLLKANKVTSNNMPTNKTEF 916
Qy 901 QLELPVKYAVYVTVSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSLSPLSLVFLVPV 960
Db 917 QLELPVKYAVYVTVSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSLSPLSLVFLVPV 976
Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDDIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDDIP 1036
Qy 1021 PFGIOEFNATLGNLSFQWYIKTSNNHLLIYSTABILFNDVSFTLLPCGGAPVRSOTET 1080
Db 1037 PFGIOEFNATLGNLSFQWYIKTSNNHLLIYSTABILFNDVSFTLLPCGGAPVRSOTET 1096
Qy 1081 KVEPFEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFRKQYKDMMSGGPPGABEPQ 1137
Db 1097 KVEPFEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFRKQYKDMMSGGPPGABEPQ 1153

RESULT 5

US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: van der Vieren, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-605-672-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTQENARGFGQSVVQLQGSRRVVVGAPOEIVAAQNRGSLVOCYDSTGSCBPI 60
Db 17 FNLDTENAMTQENARGFGQSVVQLQGSRRVVVGAPOEIVAAQNRGSLVOCYDSTGSCBPI 76
Qy 61 RLQVPVEAVNNSLGLSLAATTSPOLLACGPTVHCTCSENTVVKGLCFGLFSGNLQOQPOK 120
Db 77 RLQVPVEAVNNSLGLSLAATTSPOLLACGPTVHCTCSENTVVKGLCFGLFSGNLQOQPOK 136
Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIIPDFFRMKEFVSTVMEQKKSKTLPSLMQYSBEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIIPDFFRMKEFVSTVMEQKKSKTLPSLMQYSBEF 196
Qy 181 RIHFTFKFQONPNRSLVKGITQLLGRTHATATGVRKVIKRLNLTNGARKNAKILLVI 240
Db 197 RIHFTFKFQONPNRSLVKGITQLLGRTHATATGVRKVIKRLNLTNGARKNAKILLVI 256
Qy 241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQELNLTIAASKPRDHVQVN 300
Db 257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQELNLTIAASKPRDHVQVN 316
Qy 301 NPEALKTIQNLREKIFAIETGTQSSSSFEHEMSQBGFSAAITNGPILLSVTGSDWAG 360
Db 317 NPEALKTIQNLREKIFAIETGTQSSSSFEHEMSQBGFSAAITNGPILLSVTGSDWAG 376
Qy 361 GVPYLSKSTKSTFINNTRVDSMDNDAYLGAAAILRNKRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVPYLSKSTKSTFINNTRVDSMDNDAYLGAAAILRNKRVQSLVGLGAPRYOHIGLVAMFR 436
Qy 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYETRGQVSVCP 480
Db 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYETRGQVSVCP 496
Qy 481 PRGQARWQCDAYLGEQOPKRGAAITVLGDVNGDKLTDVAIGAPEEDNRGAVL 540
Db 497 PRGQARWQCDAYLGEQOPKRGAAITVLGDVNGDKLTDVAIGAPEEDNRGAVL 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFQSSLSGGQDLTMDGLVLTGAGQHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFQSSLSGGQDLTMDGLVLTGAGQHVLLRSQ 616
Qy 601 PVLRYKAIMBEPNPREVARNVFECDQVVKGEAGVRVCLHVOKSTRDLRSGQIQSVVT 660
Db 617 PVLRYKAIMBEPNPREVARNVFECDQVVKGEAGVRVCLHVOKSTRDLRSGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVFNENKSTRTOVLGLTQTCETLKLQLPNCIBDPVSPVILRLNF 720
Db 677 YDLALDSGRPHSRVFNENKSTRTOVLGLTQTCETLKLQLPNCIBDPVSPVILRLNF 736
Qy 721 SLVGTPLSAFNLRPVLAEDAORLFTALPPPEKNGCNDNICDDLSITSPFMSLCLVVG 780
Db 737 SLVGTPLSAFNLRPVLAEDAORLFTALPPPEKNGCNDNICDDLSITSPFMSLCLVVG 796
Qy 781 GPREFNTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORSQBSWELACESASSTEV 840
Db 797 GPREFNTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORSQBSWELACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGNKLLKANKVTSNNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGNKLLKANKVTSNNMPTNKTEF 916
Qy 901 QLELPVKYAVYVTVSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSLSPLSLVFLVPV 960
Db 917 QLELPVKYAVYVTVSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSLSPLSLVFLVPV 976
Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDDIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDDIP 1036

QY 1021 PFGIOEFNATLKNLSFDWYIKTSNHLIIVSTABILFNDVSFTLLPQCGAFVRSOTET 1080
DB 1037 PFGIOEFNATLKNLSFDWYIKTSNHLIIVSTABILFNDVSFTLLPQCGAFVRSOTET 1096
QY 1081 KVEPFVEPNPLPIVGVSSVGLILLALITAALYKLGFFKQYKDMMSGGPPGABEQ 1137
DB 1097 KVEPFVEPNPLPIVGVSSVGLILLALITAALYKLGFFKQYKDMMSGGPPGABEQ 1153

RESULT 6
US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-482-293A-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTQENARGFGQSVWLOGSRVVGAPQIIVANQSGSYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTQENARGFGQSVWLOGSRVVGAPQIIVANQSGSYQCDYSTGSCPEI 76

QY 61 RLQPVFVAVNMSLGLSLAATTSPTQLACGPTVHQTCSNTYVKGACFLFGSNLRQOPQK 120
DB 77 RLQPVFVAVNMSLGLSLAATTSPTQLACGPTVHQTCSNTYVKGACFLFGSNLRQOPQK 136

QY 121 PFALRGCPQEDSDIAFLIDSGSIIIPDFRMKEFVSTWELKSKTFLSLMQYSEEP 180

DB 137 PFALRGCPQEDSDIAFLIDSGSIIIPDFRMKEFVSTWELKSKTFLSLMQYSEEP 196
QY 181 RIHFTFKFQNNPNRSLVKPITOLLGRTHATGVRKIVIRELLNITNGARKNAFKILIVI 240
DB 197 RIHFTFKFQNNPNRSLVKPITOLLGRTHATGVRKIVIRELLNITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAPFRSEKSRQELNTIASKPPRDRHVQFN 300
DB 257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAPFRSEKSRQELNTIASKPPRDRHVQFN 316
QY 301 NFEALKTIONLREKIFAIEGTQTGSSSFHEMSQEGFSAAITNGHLLSTVYGSYDWAG 360
DB 317 NFEALKTIONLREKIFAIEGTQTGSSSFHEMSQEGFSAAITNGHLLSTVYGSYDWAG 376
QY 361 GVPLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVLGARYQHIGLVAMFR 420
DB 377 GVPLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVLGARYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDVLVIGAPHYETRGQGVSVCLP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDVLVIGAPHYETRGQGVSVCLP 496
QY 481 PRGQARWQCDAYLYGEOGQPMWREGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRGQARWQCDAYLYGEOGQPMWREGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTWGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTWGAQGHVLLRSQ 616
QY 601 PVLKVAIMEFNPREVARNVFCNDQVVKGEAGVRVCLHVOKSTRDLRREGQIQSVVT 660
DB 617 PVLKVAIMEFNPREVARNVFCNDQVVKGEAGVRVCLHVOKSTRDLRREGQIQSVVT 676
QY 661 YDLALDSGRPHSRAVFNETKNSTRQTQVLGHTQTCETLKLQLPNCIEDPVPVILRLNF 720
DB 677 YDLALDSGRPHSRAVFNETKNSTRQTQVLGHTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAORLFTALPPFEKNCNDNICODDLSITFSMSLDCLVVG 780
DB 737 SLVGTPLSAFGLNRPVLAEDAORLFTALPPFEKNCNDNICODDLSITFSMSLDCLVVG 796
QY 781 GPREFNVTVTVRNDGEDSVYRTQVTFPPFDLSYRKVSTLQNRQSRMSRLACESASSTEV 840
DB 797 GPREFNVTVTVRNDGEDSVYRTQVTFPPFDLSYRKVSTLQNRQSRMSRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANYTSENNPRTNKTFE 900
DB 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANYTSENNPRTNKTFE 916
QY 901 QLELPVKYAVVMVVTSHGVSTKYLNFASNTSRVMQHOYQVSNLQORSLPISLVFLVPV 960
DB 917 QLELPVKYAVVMVVTSHGVSTKYLNFASNTSRVMQHOYQVSNLQORSLPISLVFLVPV 976
QY 961 RLNQTVIWDPRQVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDIP 1020
DB 977 RLNQTVIWDPRQVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDIP 1036
QY 1021 PFGIOEFNATLKNLSFDWYIKTSNHLIIVSTABILFNDVSFTLLPQCGAFVRSOTET 1080
DB 1037 PFGIOEFNATLKNLSFDWYIKTSNHLIIVSTABILFNDVSFTLLPQCGAFVRSOTET 1096
QY 1081 KVEPFVEPNPLPIVGVSSVGLILLALITAALYKLGFFKQYKDMMSGGPPGABEQ 1137
DB 1097 KVEPFVEPNPLPIVGVSSVGLILLALITAALYKLGFFKQYKDMMSGGPPGABEQ 1153

RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
City: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,363
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-943-363-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAANORGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAANORGSLYQCDYSTGSCPEI 76
QY 61 RLOVPVAVNMVSLGLSLAATTSPPQLACGPTVHQCSENTYVKGCLFLFGSNLRQOPQK 120
DB 77 RLOVPVAVNMVSLGLSLAATTSPPQLACGPTVHQCSENTYVKGCLFLFGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLDGSIIIPDPRMKEFVSTVMEQLKSKTLFSLMQSYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLDGSIIIPDPRMKEFVSTVMEQLKSKTLFSLMQSYSEEF 196
QY 181 RIHFTKFEFQNNPRLSLVKPIITOLGRTHATGVRKVIKRELLNITNGARKNAFKILVI 240
DB 197 RIHFTKFEFQNNPRLSLVKPIITOLGRTHATGVRKVIKRELLNITNGARKNAFKILVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRVVGDAFRSEKSRQBLNTIASKPPRDHVFQVN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRVVGDAFRSEKSRQBLNTIASKPPRDHVFQVN 316
QY 301 NFALKTIQNLQREKIFAISGTQYTGSSSSPEHEMSQEGPSAAITNGPILSTVGSYDWAG 360
DB 317 NFALKTIQNLQREKIFAISGTQYTGSSSSPEHEMSQEGPSAAITNGPILSTVGSYDWAG 376

QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGVAAAILLRNVSQSLVGLGAPRYOHIGLIVAMFR 420
DB 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGVAAAILLRNVSQSLVGLGAPRYOHIGLIVAMFR 436
QY 421 QNTGMWESNANVKCTQIGAYFGASLCSVDVDSNGSSTDLVILGAPHYHTEOTGGGVSVCP 480
DB 437 QNTGMWESNANVKCTQIGAYFGASLCSVDVDSNGSSTDLVILGAPHYHTEOTGGGVSVCP 496
QY 481 PRGORARWQCDVLYGSGOGPWGRFGAALTIVLGVNDGDKLTDVAIGAPGEEDNRGAVL 540
DB 497 PRGORARWQCDVLYGSGOGPWGRFGAALTIVLGVNDGDKLTDVAIGAPGEEDNRGAVL 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLLSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLLSQ 616
QY 601 PVLRVKAIMENPREVARNVFECDNDVVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLRVKAIMENPREVARNVFECDNDVVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSTRRTQVGLGTCTETLKLQLPNCIEDDPVPIVLRNF 720
DB 677 YDLALDSGRPHSRVAFNETKNSTRRTQVGLGTCTETLKLQLPNCIEDDPVPIVLRNF 736
QY 721 SLVGTPLSAGFNLRPVLAEADQRLFTALFPPEKKNCGNDNICODDLISITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAGFNLRPVLAEADQRLFTALFPPEKKNCGNDNICODDLISITFSFMSLDCLVVG 796
QY 781 GPREFNVTIVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRVRLACESASSTEV 840
DB 797 GPREFNVTIVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRVRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVTNITPDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 857 SGALKSTSCSINHPIPPENSEVTNITPDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPKYAVVAVVVTSHGVSTKYLNFTASENTRVMOHOYOVSNLQSRSLPISLVFLVPV 960
DB 917 QLELPKYAVVAVVVTSHGVSTKYLNFTASENTRVMOHOYOVSNLQSRSLPISLVFLVPV 976
QY 961 RLNQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLABELKAPVAVVNCISLAVCORIQCDIP 1020
DB 977 RLNQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLABELKAPVAVVNCISLAVCORIQCDIP 1036
QY 1021 FFGIOEEFNATLKGMLSPDWYIKTSHNHLIVSTAEILFNDVSFTLLPQOGAFVRSOTET 1080
DB 1037 FFGIOEEFNATLKGMLSPDWYIKTSHNHLIVSTAEILFNDVSFTLLPQOGAFVRSOTET 1096
QY 1081 KVEPEVNPPLIIVGSSVGGILLALITALYKLGFFKQYKQVKNVSGGPPGABPQ 1137
DB 1097 KVEPEVNPPLIIVGSSVGGILLALITALYKLGFFKQYKQVKNVSGGPPGABPQ 1153

RESULT 8

US-09-193-043-3
Sequence 3, Application US/09193043
Patent No. 6251395
GENERAL INFORMATION:
APPLICANT: Gallatin, Michael W.
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 6251395e1 Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/193,043
CURRENT FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: 08/173,497
EARLIER FILING DATE: 1993-12-23
EARLIER APPLICATION NUMBER: 08/286,889
EARLIER FILING DATE: 1994-08-05
EARLIER APPLICATION NUMBER: 08/362,652
EARLIER FILING DATE: 1994-12-21
EARLIER APPLICATION NUMBER: 08/943,363
EARLIER FILING DATE: 1997-10-03

```

; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-193-043-3

Query Match      99.9%; Score 5868; DB 3; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARFGQSVVQLQGSRRVVGAPQEIIVAAANQRGSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFOENARFGQSVVQLQGSRRVVGAPQEIIVAAANQRGSLYQCDYSTGSCBPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCSNTYVKGCLFPGSNLRQPOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCSNTYVKGCLFPGSNLRQPOK 136
QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKBFVSTVMEQLKSKTLFSLMOYSEEF 180
DB 137 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKBFVSTVMEQLKSKTLFSLMOYSEEF 196
QY 181 RIHFTFKFQNNPNRSLVKPITQLGRTHATGVRKVIPELLNITNGARKNAFKILIVI 240
DB 197 RIHFTFKFQNNPNRSLVKPITQLGRTHATGVRKVIPELLNITNGARKNAFKILIVI 256
QY 241 TDGEKFGPGLGYEDVTPADREGVIRYVIGVDGDFSEKSRQELNTIASKPRDHVFN 300
DB 257 TDGEKFGPGLGYEDVTPADREGVIRYVIGVDGDFSEKSRQELNTIASKPRDHVFN 316
QY 301 NFEALKTQNLREKIPAEIGTQTGSSSFHEMSQEGFSAAITSNGLPILSTVGSYDNAG 360
DB 317 NFEALKTQNLREKIPAEIGTQTGSSSFHEMSQEGFSAAITSNGLPILSTVGSYDNAG 376
QY 361 GYFLYTSKESKPIFNTRVDSNDMDAYLAGVAAAILLRNVQSVILGAPYCHIGLVAMFR 420
DB 377 GYFLYTSKESKPIFNTRVDSNDMDAYLAGVAAAILLRNVQSVILGAPYCHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVCP 496
QY 481 PRQORARWQCDVLYGEQOPWGRFGAALTVDGVDGDKLTVAIGAPGEEDNRGAVYLF 540
DB 497 PRQORARWQCDVLYGEQOPWGRFGAALTVDGVDGDKLTVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGQAQGHVILLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGQAQGHVILLRSQ 616
QY 601 PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRQGTQSVVT 660
DB 617 PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRQGTQSVVT 676
QY 661 YDLALDSGRPHSAVNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720
DB 677 YDLALDSGRPHSAVNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 736
QY 721 SLVGTPLSAFGRNRPVLAEDAQRLFTALPFPFKNCNDNMCDDLSITFSPMSLDCLVVG 780
DB 737 SLVGTPLSAFGRNRPVLAEDAQRLFTALPFPFKNCNDNMCDDLSITFSPMSLDCLVVG 796
QY 781 GPREFNVTVTRNDGDSYRTQTFEPFLDLSYRKVSTLQNSORSWRILACESASSTEV 840
DB 797 GPREFNVTVTRNDGDSYRTQTFEPFLDLSYRKVSTLQNSORSWRILACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNPRTNKTF 900
DB 857 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNPRTNKTF 916
QY 901 QLELPVKYAVYVMVTSHGVS TKYLNFTASENTSRVMQHQVQVSNLQORSIPISLVFLVPV 976
DB 961 RLNQTVIWDPRQVTPFSENLSASTCHTKERLPSHSDFLAELKAPVYVNCISIAVCORIQCDIP 1020
DB 977 RLNQTVIWDPRQVTPFSENLSASTCHTKERLPSHSDFLAELKAPVYVNCISIAVCORIQCDIP 1036
QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPFGQAFVRSQET 1080
DB 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPFGQAFVRSQET 1096
QY 1081 KVBEPFEVFNPLPIVGVSSVGLLILLALITAAALYKLGFFKQYKDMMEGEGPFGAEPQ 1137
DB 1097 KVBEPFEVFNPLPIVGVSSVGLLILLALITAAALYKLGFFKQYKDMMEGEGPFGAEPQ 1153

RESULT 9
US-09-688-307A-3
; Sequence 3, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6432404el Human Beta-2
; FILE REFERENCE: 27866/36846
; CURRENT APPLICATION NUMBER: US/09/688, 307A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/193, 043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605, 672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173, 497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286, 889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362, 652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943, 363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-307A-3

Query Match      99.9%; Score 5868; DB 4; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARFGQSVVQLQGSRRVVGAPQEIIVAAANQRGSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFOENARFGQSVVQLQGSRRVVGAPQEIIVAAANQRGSLYQCDYSTGSCBPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCSNTYVKGCLFPGSNLRQPOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCSNTYVKGCLFPGSNLRQPOK 136
QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKBFVSTVMEQLKSKTLFSLMOYSEEF 180
DB 137 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKBFVSTVMEQLKSKTLFSLMOYSEEF 196
QY 181 RIHFTFKFQNNPNRSLVKPITQLGRTHATGVRKVIPELLNITNGARKNAFKILIVI 240
DB 197 RIHFTFKFQNNPNRSLVKPITQLGRTHATGVRKVIPELLNITNGARKNAFKILIVI 256
QY 241 TDGEKFGPGLGYEDVTPADREGVIRYVIGVDGDFSEKSRQELNTIASKPRDHVFN 300
DB 257 TDGEKFGPGLGYEDVTPADREGVIRYVIGVDGDFSEKSRQELNTIASKPRDHVFN 316
QY 301 NFEALKTQNLREKIPAEIGTQTGSSSFHEMSQEGFSAAITSNGLPILSTVGSYDNAG 360
DB 317 NFEALKTQNLREKIPAEIGTQTGSSSFHEMSQEGFSAAITSNGLPILSTVGSYDNAG 376
QY 361 GYFLYTSKESKPIFNTRVDSNDMDAYLAGVAAAILLRNVQSVILGAPYCHIGLVAMFR 420
DB 377 GYFLYTSKESKPIFNTRVDSNDMDAYLAGVAAAILLRNVQSVILGAPYCHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVCP 496
QY 481 PRQORARWQCDVLYGEQOPWGRFGAALTVDGVDGDKLTVAIGAPGEEDNRGAVYLF 540
DB 497 PRQORARWQCDVLYGEQOPWGRFGAALTVDGVDGDKLTVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGQAQGHVILLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGQAQGHVILLRSQ 616
QY 601 PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRQGTQSVVT 660
DB 617 PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRQGTQSVVT 676
QY 661 YDLALDSGRPHSAVNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720
DB 677 YDLALDSGRPHSAVNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 736
QY 721 SLVGTPLSAFGRNRPVLAEDAQRLFTALPFPFKNCNDNMCDDLSITFSPMSLDCLVVG 780
DB 737 SLVGTPLSAFGRNRPVLAEDAQRLFTALPFPFKNCNDNMCDDLSITFSPMSLDCLVVG 796
QY 781 GPREFNVTVTRNDGDSYRTQTFEPFLDLSYRKVSTLQNSORSWRILACESASSTEV 840
DB 797 GPREFNVTVTRNDGDSYRTQTFEPFLDLSYRKVSTLQNSORSWRILACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNPRTNKTF 900
DB 857 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNPRTNKTF 916
QY 901 QLELPVKYAVYVMVTSHGVS TKYLNFTASENTSRVMQHQVQVSNLQORSIPISLVFLVPV 976
DB 961 RLNQTVIWDPRQVTPFSENLSASTCHTKERLPSHSDFLAELKAPVYVNCISIAVCORIQCDIP 1020
DB 977 RLNQTVIWDPRQVTPFSENLSASTCHTKERLPSHSDFLAELKAPVYVNCISIAVCORIQCDIP 1036
QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPFGQAFVRSQET 1080
DB 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPFGQAFVRSQET 1096
QY 1081 KVBEPFEVFNPLPIVGVSSVGLLILLALITAAALYKLGFFKQYKDMMEGEGPFGAEPQ 1137
DB 1097 KVBEPFEVFNPLPIVGVSSVGLLILLALITAAALYKLGFFKQYKDMMEGEGPFGAEPQ 1153
```

317 NFEALKTQNLREKIFALEGTQSSSSFEHMSQEGPSAITSNGPILLSVTGSDYDAG 376
361 GVFLYTSKESKSTFINNTRVDSMDNDAYLGYAAAILRNVRQSLVGLGAPYQHIGLVAMR 420
377 GVFLYTSKESKSTFINNTRVDSMDNDAYLGYAAAILRNVRQSLVGLGAPYQHIGLVAMR 436
421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGGQSVCP 480
437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGGQSVCP 496
481 PRGORARWQCDVLYGEOQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
497 PRGORARWQCDVLYGEOQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 600
557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 616
601 PVLRVKALMEFNPVARNVFECDVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT 660
617 PVLRVKALMEFNPVARNVFECDVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT 676
661 YDLALDSGRPHSAFNPVLAEDAOBLFTALPFEKNCNDNICQDDLSTITFSMGLDCLVVG 720
677 YDLALDSGRPHSAFNPVLAEDAOBLFTALPFEKNCNDNICQDDLSTITFSMGLDCLVVG 736
721 SLVGTPLSAFNGNRPVLAEDAOBLFTALPFEKNCNDNICQDDLSTITFSMGLDCLVVG 780
737 SLVGTPLSAFNGNRPVLAEDAOBLFTALPFEKNCNDNICQDDLSTITFSMGLDCLVVG 796
781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACSSASSTEV 840
797 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACSSASSTEV 856
841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKXANTYSENMMERTNKTEF 900
857 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKXANTYSENMMERTNKTEF 916
901 QLELPVKYAVYVTVSHGVSTKYNFTASNTSRVMOHQVQVSNLQORSIPISLVFLVPV 960
917 QLELPVKYAVYVTVSHGVSTKYNFTASNTSRVMOHQVQVSNLQORSIPISLVFLVPV 976
961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAQVWNCISAVCQRIQCDIP 1020
977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAQVWNCISAVCQRIQCDIP 1036
1021 PFCIOBEPNATLKNLSFDVYKTSNNHLLIVSTAELLFNDVSVFTLLPGOGAFVRSQTE 1080
1037 PFCIOBEPNATLKNLSFDVYKTSNNHLLIVSTAELLFNDVSVFTLLPGOGAFVRSQTE 1086
1081 KVEPFEVNPPLTVGSSVGLLALLITALYKLGFFKQYKXDMMSGEGPPGABPQ 1137
1097 KVEPFEVNPPLTVGSSVGLLALLITALYKLGFFKQYKXDMMSGEGPPGABPQ 1153

RESULT 10

US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. 6620915
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. 6620915el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350.259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193.043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173.497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286.889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362.652

; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943.363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.9%; Score 5868; DB 4; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTQENARGFGQSVVQLQGSRRVVGAPQBEIVAAANQORSLVQCDYSTGSCSCEPI 60
DB 17 FNLDTENAMTQENARGFGQSVVQLQGSRRVVGAPQBEIVAAANQORSLVQCDYSTGSCSCEPI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTYVVKGLCLFLGNSLQQPOK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTYVVKGLCLFLGNSLQQPOK 136
QY 121 FPEALRGCPQEDSDIAPLIDGSGSIIPHDFRMKKEFVSTVMEOLKSKTLPISLMQYSBEF 180
DB 137 FPEALRGCPQEDSDIAPLIDGSGSIIPHDFRMKKEFVSTVMEOLKSKTLPISLMQYSBEF 196
QY 181 RHFTFKFQFNNPNSRLVKPITQLLGRTHATGVRKVIRELNITNGARKNAFKILYI 240
DB 197 RHFTFKFQFNNPNSRLVKPITQLLGRTHATGVRKVIRELNITNGARKNAFKILYI 256
QY 241 TDGEXFGDPLGYEDVIPEADREGVIRYVIGVGDAPFRSEKSRQELNITLASKPRDHVQFN 300
DB 257 TDGEXFGDPLGYEDVIPEADREGVIRYVIGVGDAPFRSEKSRQELNITLASKPRDHVQFN 316
QY 301 NFEALKTQNLREKIFALEGTQSSSSFEHMSQEGPSAITSNGPILLSVTGSDYDAG 360
DB 317 NFEALKTQNLREKIFALEGTQSSSSFEHMSQEGPSAITSNGPILLSVTGSDYDAG 376
QY 361 GVFLYTSKESKSTFINNTRVDSMDNDAYLGYAAAILRNVRQSLVGLGAPYQHIGLVAMR 420
DB 377 GVFLYTSKESKSTFINNTRVDSMDNDAYLGYAAAILRNVRQSLVGLGAPYQHIGLVAMR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGGQSVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGGQSVCP 496
QY 481 PRGORARWQCDVLYGEOQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRGORARWQCDVLYGEOQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 616
QY 601 PVLRVKALMEFNPVARNVFECDVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT 660
DB 617 PVLRVKALMEFNPVARNVFECDVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT 676
QY 661 YDLALDSGRPHSAFNPVLAEDAOBLFTALPFEKNCNDNICQDDLSTITFSMGLDCLVVG 720
DB 677 YDLALDSGRPHSAFNPVLAEDAOBLFTALPFEKNCNDNICQDDLSTITFSMGLDCLVVG 736
QY 721 SLVGTPLSAFNGNRPVLAEDAOBLFTALPFEKNCNDNICQDDLSTITFSMGLDCLVVG 780
DB 737 SLVGTPLSAFNGNRPVLAEDAOBLFTALPFEKNCNDNICQDDLSTITFSMGLDCLVVG 796
QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACSSASSTEV 840
DB 797 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACSSASSTEV 856
QY 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKXANTYSENMMERTNKTEF 900

857 SGALKSTCSINHPDPFENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEP 916
 901 QLELPVKYAVYVWVSHGYSKYLNFTASENISRVMQHQYQVSNLQORSLSPLSLVPLVPV 960
 917 QLELPVKYAVYVWVSHGYSKYLNFTASENISRVMQHQYQVSNLQORSLSPLSLVPLVPV 976
 961 RLNQTVWDRPQVTESENLSSTCHTKERLPSSHDSFLAELKAPVWNCISAVCQRIQCDIP 1020
 977 RLNQTVWDRPQVTESENLSSTCHTKERLPSSHDSFLAELKAPVWNCISAVCQRIQCDIP 1036
 1021 FFGIOEFNATLKGNLSFDWYIKTSHNHLIIIVSTAEILFNDSVFTLLPGQAFVRSQTEP 1080
 1037 FFGIOEFNATLKGNLSFDWYIKTSHNHLIIIVSTAEILFNDSVFTLLPGQAFVRSQTEP 1096
 1081 KVEPFEVNPFLPIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1137
 1097 KVEPFEVNPFLPIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1153

RESULT 11

US-08-476-062A-43
 ; Sequence 43, Application US/08476062A
 ; Patent No. 587275
 ; GENERAL INFORMATION:
 ; APPLICANT: Attnaut, M. Amin
 ; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
 ; NUMBER OF SEQUENCES: 53
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: US
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows95
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/476,062A
 ; FILING DATE: 07-JUN-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/216,081
 ; FILING DATE: 21-MAR-1994
 ; APPLICATION NUMBER: 07/537,830
 ; FILING DATE: 04-JAN-1991
 ; APPLICATION NUMBER: 07/539,842
 ; FILING DATE: 18-JUN-1990
 ; APPLICATION NUMBER: 07/212,573
 ; FILING DATE: 28-JUN-1988
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Freeman, John W.
 ; REGISTRATION NUMBER: 29,066
 ; REFERENCE/DOCKET NUMBER: 00786/068003
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 43:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1152 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; FRAGMENT TYPE: internal
 ; US-08-476-062A-43

Query Match 99.4%; Score 5837.5; DB 2; Length 1152;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1130; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

1 FNLDTENAMTFOENARGFGOSVWLOGSRVVGAPQEIIVAAANORGSLYQCDYSTGSCBPI 60
 17 FNLDTENAMTFOENARGFGOSVWLOGSRVVGAPQEIIVAAANORGSLYQCDYSTGSCBPI 76
 61 RLQVPVEAVNMSLGLSLAATTPPQOLLACGPTVHTCSENTYVKGCLCFPGSNLRQOQOK 120
 77 RLQVPVEAVNMSLGLSLAATTPPQOLLACGPTVHTCSENTYVKGCLCFPGSNLRQOQOK 136
 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQKKSKTLFSLMOYSSEF 180
 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQKKSKTLFSLMOYSSEF 196
 181 RHFTFKFQONNPNRSLVKDITOLLGHTHTATGVRKVIIRLLNITNGARKNAKILIVI 240
 197 RHFTFKFQONNPNRSLVKDITOLLGHTHTATGVRKVIIRLLNITNGARKNAKILIVI 256
 241 TDGEKFGDPLGVEDVIPEADREGVIRYVIGVDGAFRSEKSEQLNLTASKPDRDHVQVN 300
 257 TDGEKFGDPLGVEDVIPEADREGVIRYVIGVDGAFRSEKSEQLNLTASKPDRDHVQVN 316
 301 NFEALKTIONOLREKI PAIEGTOTGSSSSFEHMSQBSAAITNSGPLELSTVGSYDWAQ 360
 317 NFEALKTIONOLREKI PAIEGTOTGSSSSFEHMSQBSAAITNSGPLELSTVGSYDWAQ 376
 361 GVFLYTSKEKSTFINMTVDSMDNDAYLGAAAAIILNRNVQSLVGLGAPRYOHIGLVAWER 420
 377 GVFLYTSKEKSTFINMTVDSMDNDAYLGAAAAIILNRNVQSLVGLGAPRYOHIGLVAWER 436
 421 QNTGMWESNANVKGTGAYFGASLCSVDVDSNGSTDLVIGAPHYHYOTRGQVSVCP 480
 437 QNTGMWESNANVKGTGAYFGASLCSVDVDSNGSTDLVIGAPHYHYOTRGQVSVCP 496
 481 FRGQRRWQCDVLYGSGQGWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLP 540
 497 FRGQRRWQCDVLYGSGQGWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLP 555
 541 HGTSGGSPSHSRIAGSKLSPLOYGOSLSGGDLTMDGLVDLTVGAGCHVLLRSQ 600
 556 HGTSGGSPSHSRIAGSKLSPLOYGOSLSGGDLTMDGLVDLTVGAGCHVLLRSQ 615
 601 PVLRYKALMEFNPREVARNVFECDQVWKGKAGEVAVCLHVQKSTRDRLRREGIQSVVT 660
 616 PVLRYKALMEFNPREVARNVFECDQVWKGKAGEVAVCLHVQKSTRDRLRREGIQSVVT 675
 661 YDLALDSGRPHSRAVFNETKNSITROTIVLCTCTETLKLQPNCTEDPVSPIVLELNF 720
 676 YDLALDSGRPHSRAVFNETKNSITROTIVLCTCTETLKLQPNCTEDPVSPIVLELNF 735
 721 SLVGTPLSAFGLNRPVLAEDAQRLEFALFPFKKNCNDNICQDDLSITFSFMSLDCLVVG 780
 736 SLVGTPLSAFGLNRPVLAEDAQRLEFALFPFKKNCNDNICQDDLSITFSFMSLDCLVVG 795
 781 GPREFNVTTVVRNDGEDSVRTOTFPPLDLSYRKVSTLONORSORSWRLACSSASSTEV 840
 796 GPREFNVTTVVRNDGEDSVRTOTFPPLDLSYRKVSTLONORSORSWRLACSSASSTEV 855
 841 SGALKSTCSINHPDPFENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEP 900
 856 SGALKSTCSINHPDPFENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEP 915
 901 QLELPVKYAVYVWVSHGYSKYLNFTASENISRVMQHQYQVSNLQORSLSPLSLVPLVPV 960
 916 QLELPVKYAVYVWVSHGYSKYLNFTASENISRVMQHQYQVSNLQORSLSPLSLVPLVPV 975
 961 RLNQTVWDRPQVTESENLSSTCHTKERLPSSHDSFLAELKAPVWNCISAVCQRIQCDIP 1020
 976 RLNQTVWDRPQVTESENLSSTCHTKERLPSSHDSFLAELKAPVWNCISAVCQRIQCDIP 1035
 1021 FFGIOEFNATLKGNLSFDWYIKTSHNHLIIIVSTAEILFNDSVFTLLPGQAFVRSQTEP 1080
 1036 FFGIOEFNATLKGNLSFDWYIKTSHNHLIIIVSTAEILFNDSVFTLLPGQAFVRSQTEP 1095
 1081 KVEPFEVNPFLPIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1137

Db 1096 KVEPEVFNPLPLVGVSSVGGILLALITAAALYKLGFFKQYKMMSEGGPGABPQ 1152

RESULT 12
PCT-US96-01314-43
Sequence 43, Application PC/TUS9601314
GENERAL INFORMATION:
APPLICANT: M. Amin Arnaout
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
TITLE OF INVENTION: ANTAGONISTS
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1152
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US96-01314-43

Query Match 99.4%; Score 5837.5; DB 5; Length 1152;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1130; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAPQEI VAAANQSGSLYQCDYSTGSCBPI 60
Db 17 FNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAPQEI VAAANQSGSLYQCDYSTGSCBPI 76

Qy 61 RLQVPVEANMSIGLSLAATSPQQLACGPTVHQTCSNTYVKGICFLPGSLNLCQQPK 120
Db 77 RLQVPVEANMSIGLSLAATSPQQLACGPTVHQTCSNTYVKGICFLPGSLNLCQQPK 136

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKKSKTLFSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKKSKTLFSLMOYSEEF 196

Qy 181 RHPTFKFQNNPNRSLVKPTQLLGRTHATGKVRKVIKRELLNTNGARKNAFKILIVI 240
Db 197 RHPTFKFQNNPNRSLVKPTQLLGRTHATGKVRKVIKRELLNTNGARKNAFKILIVI 256

Qy 241 TGEKFGDPLGYEDVTPEDRGEVIRYVIGVDAPRSEKSRQELNTIASKPPRDHVFQV 300
Db 257 TGEKFGDPLGYEDVTPEDRGEVIRYVIGVDAPRSEKSRQELNTIASKPPRDHVFQV 316

Qy 301 NFEALKTIONQAREKIFAIEGTQTSSSSFEHMSQBGPSAAITNGPILLSTVGVSDWAG 360
Db 317 NFEALKTIONQAREKIFAIEGTQTSSSSFEHMSQBGPSAAITNGPILLSTVGVSDWAG 376

Qy 361 GVFLYTSKEKSTFTINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFTINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMPR 436

Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQSVQVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQSVQVCP 496

Qy 481 PRGQARWQCDVLYGEGQPGWRFGAALTIVLGDVNGDKLTDVAIGAPBEDNRGAVYLF 540
Db 497 PRG-RARWQCDVLYGEGQPGWRFGAALTIVLGDVNGDKLTDVAIGAPBEDNRGAVYLF 555

Qy 541 HGTSGSISPSHSORLQAGSLKSLPRLOYFGQSLSGQDLTMDGLVLTVGAQGHVLLRSQ 600
Db 556 HGTSGSISPSHSORLQAGSLKSLPRLOYFGQSLSGQDLTMDGLVLTVGAQGHVLLRSQ 615

Qy 601 PVLRVKAIMBENPREVARNVFECNDQVVGKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 616 PVLRVKAIMBENPREVARNVFECNDQVVGKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 675

Qy 661 YDLALDSGRPHSRVAFNETKSTRQTVLGLTOTCETLKLQNCIEDPVPVILRLNF 720
Db 676 YDLALDSGRPHSRVAFNETKSTRQTVLGLTOTCETLKLQNCIEDPVPVILRLNF 735

Qy 721 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKMGNDNI CQDDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKMGNDNI CQDDLSITFSFMSLDCLVVG 795

Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRSGRSLACESASSTEV 840
Db 796 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRSGRSLACESASSTEV 855

Qy 841 SGALKSTSCSINHPIFPENSESVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSESVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 915

Qy 901 QLELPVKYAVVWVTSVSHGVSTKYLNFTASENTRSRVMOHQYQVSNLQSRSLPISLVFLVPV 960
Db 916 QLELPVKYAVVWVTSVSHGVSTKYLNFTASENTRSRVMOHQYQVSNLQSRSLPISLVFLVPV 975

Qy 961 RLNQTVIWDPRQVTFPSENLSSTCHTKERLPSHSDPLAELRKAPVNVNCIAVCORIQCDIP 1020
Db 976 RLNQTVIWDPRQVTFPSENLSSTCHTKERLPSHSDPLAELRKAPVNVNCIAVCORIQCDIP 1035

Qy 1021 PFGIQEENATLKGNSLSPDWYIKTSHNHLIVSTAEILFNDSPVTLPLPGQAFVRSQET 1080
Db 1036 PFGIQEENATLKGNSLSPDWYIKTSHNHLIVSTAEILFNDSPVTLPLPGQAFVRSQET 1095

Qy 1081 KVEPEVFNPLPLVGVSSVGGILLALITAAALYKLGFFKQYKMMSEGGPGABPQ 1137
Db 1096 KVEPEVFNPLPLVGVSSVGGILLALITAAALYKLGFFKQYKMMSEGGPGABPQ 1152

RESULT 13
5424399-2
Patent No. 5424399
APPLICANT: ARNAOUT, M. AMIN
TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
NUMBER OF SEQUENCES: 12
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/78,871
FILING DATE: 16-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 539,842
FILING DATE: 18-JUN-1990
APPLICATION NUMBER: 212,573
FILING DATE: 28-JUN-1988
SEQ ID NO: 2;
LENGTH: 1152
5424399-2

Query Match 99.4%; Score 5837.5; DB 6; Length 1152;

[illegible]

Search completed: June 7, 2004, 17:19:31
Job time : 22.3484 secs